

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2005, 18:20:14 ; Search time 24035 Seconds
(without alignments)
11858.273 Million cell updates/sec

Title: US-10-715-665-6

Perfect score: 5882

Sequence: 1 tcgcgcttctcgatgatgc.....tatcacgagcccttcgtc 5882

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5882	100.0	5882	6	AX154501 Sequence
2	4617.2	78.5	5459	6	AX154499 Sequence
3	4190	71.2	5128	6	AX154497 Sequence
4	3713.6	63.1	6299	6	AX164584 Sequence
5	2793	47.5	5130	12	AF286076 Expression
6	2658.4	45.2	4276	6	AX154496 Sequence
7	2601.2	44.2	4282	6	AX164583 Sequence
8	2397.8	40.8	6050	6	CQ855138 Sequence
9	2333.2	39.7	7551	6	CQ795336 Sequence
10	2333.2	39.7	7551	6	CQ867240 Sequence
11	2290.6	38.9	6577	6	AX427812 Sequence
12	2283.4	38.8	6505	6	AX427777 Sequence
13	2283	38.8	6460	6	AX427796 Sequence
14	2281.8	38.8	9620	6	AX164579 Sequence
15	2281.8	38.8	9620	6	AX164581 Sequence
16	2280.4	38.8	6577	6	AX427813 Sequence
17	2278	38.7	5411	12	CV003687
18	2275.8	38.7	6438	6	AX427919 Sequence
19	2275.8	38.7	6577	6	AX427794 Sequence

20	2270.8	38.6	6473	6	AX427862 Sequence
21	2269.2	38.6	6607	6	AX427770 Sequence
22	2266	38.5	6486	6	AX427920 Sequence
23	2263.6	38.5	6605	6	AX427867 Sequence
24	2263.6	38.5	6605	6	AX427868 Sequence
25	2262	38.5	6538	6	AX427795 Sequence
26	2261.6	38.4	6608	6	AX427884 Sequence
27	2260.6	38.4	7178	6	CQ785592 Sequence
28	2260.6	38.4	7178	6	CQ795334 Sequence
29	2260.6	38.4	7178	6	CQ867238 Sequence
30	2260	38.4	6602	6	AX427865 Sequence
31	2259.8	38.4	6609	6	AX427846 Sequence
32	2259.6	38.4	6536	6	AX427869 Sequence
33	2256.6	38.4	6505	6	AX427778 Sequence
34	2256.6	38.4	6565	6	AR235454 Sequence
35	2256	38.4	6545	6	AX427877 Sequence
36	2252	38.3	6577	6	AX427906 Sequence
37	2251.4	38.3	6466	6	AX427892 Sequence
38	2251.4	38.3	6538	6	AX427907 Sequence
39	2251.4	38.3	6586	6	AX427890 Sequence
40	2250.8	38.3	6614	6	AX427876 Sequence
41	2250.8	38.3	6617	6	AX427870 Sequence
42	2248.4	38.2	6542	6	AX427861 Sequence
43	2248.4	38.2	6610	6	AX427829 Sequence
44	2248	38.2	6623	6	AX427860 Sequence
45	2246.4	38.2	6626	6	AX427878 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX154501 5882 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6 from Patent WO0138358.
ACCESSION AX154501
VERSION AX154501.1 GI:14536111
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Selby, M.C., Glazer, E.C. and Houghton, M.C.
TITLE Hbv/hcv virus-like particle
JOURNAL Patent: WO 0138358-A 6 31-MAY-2001;
CHIRON CORPORATION (US)

FEATURES
source
i. .5882
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="plasmid pCMV-II-E2661-sag"
1992. .3587
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="Cac42602.1"
/db_xref="GI:14536112"
/translation="MDAMKRGCLCCVLLCGAVFSPASETHVTGSGAGHTVSGFVSL
LAPGAKQNVQLINTGSHLNTALNCDSLTGMLAGLFYHHKFNSSGCPKRLASCR
PLTDFDQGMGPISYANGSDQDPYCMHYPPKPGIVPAKSGVPCYCFPPSPVVGST
TRSGAPTSWGENDTDFVLLNTPPLGNWFCCTWMNSTGFTKVCAGPPCVTGGAGN
NLTGPTDCFRKHPDPTVSRCSGPWITPRCLVDVYRLLWHPYCTINVTIKFIMVVG
GYEHLERACNWTGRCDEDRDSEIDMENTSGFLGPLLVLAGFLLFTLITIP
QSLDSWTSNLNFGSPVCLGQNSQSPITSNHSPTSCPPICPGYRNWCLRRFLITFIL
LILCLFLVLLDYOQMLPVCPLIPGSTITSTGTCTCTTPAQGNMFPSCCCCTKPTDG
NCTCPIFSSWAFAPKYLWNASVRSWLSLLVPFQVQFVGLSPVTWLSAIIWMWYNGP
SLYSIVSPRIPLLPFFCLWYI"

CDS

ORIGIN

Query Match 100.0%; Score 5882; DB 6; Length 5882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TCGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGCACTCCCGAGACGGTCA	60
Db	1	TCGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGCACTCCCGAGACGGTCA	60
Qy	61	CAGCTTGTCTGTAAAGCGATGCCGGGAGCAGACAAGCCGTCAGGCGCGTCACGGGGTG	120
Db	61	CAGCTTGTCTGTAAAGCGATGCCGGGAGCAGACAAGCCGTCAGGCGCGTCACGGGGTG	120
Qy	121	TTGCGGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC	180
Db	121	TTGCGGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC	180
Qy	181	ACCATATGAAGCTTTTTCGAAAAGCCTAGGCGCTCCAAAAAGCCCTCTCACTTCTCTGG	240
Db	181	ACCATATGAAGCTTTTTCGAAAAGCCTAGGCGCTCCAAAAAGCCCTCTCACTTCTCTGG	240
Qy	241	AATAGCTCAGAGCGCGAGCGGCGCTCTGCGCTCTGATATAATAAAAAAATTTAGTCAGCCA	300
Db	241	AATAGCTCAGAGCGCGAGCGGCGCTCTGCGCTCTGATATAATAAAAAAATTTAGTCAGCCA	300
Qy	301	TGGGGCGGAGAAATGGGCGGAATGGGCGGGAGGGAATTTATGGCTATTGGCCATTGCAT	360
Db	301	TGGGGCGGAGAAATGGGCGGAATGGGCGGGAGGGAATTTATGGCTATTGGCCATTGCAT	360
Qy	361	ACGTTGTATCTATATCATATATGTAATTTATATGGCTCATGTCCAAATATGACCGCCA	420
Db	361	ACGTTGTATCTATATCATATATGTAATTTATATGGCTCATGTCCAAATATGACCGCCA	420
Qy	421	TGTTGACATTTGATTGATTTGACCTAGTTAATTAATAGTAATCAATTTACGGGGTCATTAGTTTCAT	480
Db	421	TGTTGACATTTGATTGATTTGACCTAGTTAATTAATAGTAATCAATTTACGGGGTCATTAGTTTCAT	480
Qy	481	AGCCCATATATGGAGTTCCGGTTTACATAACTTAGGTAAATGCGCCGCGCTGGCTGACCG	540
Db	481	AGCCCATATATGGAGTTCCGGTTTACATAACTTAGGTAAATGCGCCGCGCTGGCTGACCG	540
Qy	541	CCCAACGACCCCGCCCATTGACGTCAATATGACGTATGTTCCATAGTAAGCCCAATA	600
Db	541	CCCAACGACCCCGCCCATTGACGTCAATATGACGTATGTTCCATAGTAAGCCCAATA	600
Qy	601	GGGACTTTCCATTGACGTCAATGGTGGAGTATTTACGGTAAACTGCGCCACTGGCGAGTA	660
Db	601	GGGACTTTCCATTGACGTCAATGGTGGAGTATTTACGGTAAACTGCGCCACTGGCGAGTA	660
Qy	661	CATCAAGTGATCATATGCCAAGTCCGCGCCCTATTGACGTCAATGACGTAAGTGGCCC	720
Db	661	CATCAAGTGATCATATGCCAAGTCCGCGCCCTATTGACGTCAATGACGTAAGTGGCCC	720
Qy	721	GCCTGGCATATTGCCAGTACATGACCTTACGGGACTTTCTCTACTTGGCAGTACATCTAC	780
Db	721	GCCTGGCATATTGCCAGTACATGACCTTACGGGACTTTCTCTACTTGGCAGTACATCTAC	780
Qy	781	GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGA	840
Db	781	GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGA	840
Qy	841	TAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG	900
Db	841	TAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG	900
Qy	901	TTTTTGGGACCAAAATCAACGGGACTTTCCAAAATGCTGAATAACCCGCCCTTCACG	960
Db	901	TTTTTGGGACCAAAATCAACGGGACTTTCCAAAATGCTGAATAACCCGCCCTTCACG	960
Qy	961	CAAAATGGCGGTAGGCGGTGACGGTGGGAGGTCTATATAGCAGAGCTCGTTTAGTGAAC	1020
Db	961	CAAAATGGCGGTAGGCGGTGACGGTGGGAGGTCTATATAGCAGAGCTCGTTTAGTGAAC	1020
Qy	1021	CGTCAGATTCGCTGGAGACGCCATCCACGCTGTTTTGACCTTCCATAGAGAACACCGGGAC	1080
Db	1021	CGTCAGATTCGCTGGAGACGCCATCCACGCTGTTTTGACCTTCCATAGAGAACACCGGGAC	1080

Qy	1081	CGATCCAGCCTCCGCGCCGGGAAACGGTGCATTGGBAACGGCGATTCCCCTGGCCAAAGT	1141
Db	1081	CGATCCAGCCTCCGCGCCGGGAAACGGTGCATTGGBAACGGCGATTCCCCTGGCCAAAGT	1140
Qy	1141	GAGCTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1200
Db	1141	GAGCTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1200
Qy	1201	CTGTTTTTGGCTTGGGCGCTATACACCCCGCTCCTTATGCTATAGTGATGGTATAGCT	1260
Db	1201	CTGTTTTTGGCTTGGGCGCTATACACCCCGCTCCTTATGCTATAGTGATGGTATAGCT	1260
Qy	1261	TAGCCTATAGGTGTGGGTATTGACCAATTAATGACCACTCCCCTATTGGTGACGATACTT	1320
Db	1261	TAGCCTATAGGTGTGGGTATTGACCAATTAATGACCACTCCCCTATTGGTGACGATACTT	1320
Qy	1321	TCCATTACTAATCCATAAACATGCTCTTTGGCCAACTATCTCTATTGGGTATATGCCAA	1380
Db	1321	TCCATTACTAATCCATAAACATGCTCTTTGGCCAACTATCTCTATTGGGTATATGCCAA	1380
Qy	1381	TACTCTGTCTCTCAGAGACTGACACGGACTCTGTATTTTTCACAGGATGGGGTCCATTAT	1440
Db	1381	TACTCTGTCTCTCAGAGACTGACACGGACTCTGTATTTTTCACAGGATGGGGTCCATTAT	1440
Qy	1441	TATTTACAAATTCCACATATACAAACACGCCGTCCCCTGGTCCCGCAGTTTTTATTAACA	1500
Db	1441	TATTTACAAATTCCACATATACAAACACGCCGTCCCCTGGTCCCGCAGTTTTTATTAACA	1500
Qy	1501	TAGGTGGGATCTCCGACATCTCGGTACGTGTTCGGACATGGGCTCTTCTCGGTAGC	1560
Db	1501	TAGCGTGGGATCTCCGACATCTCGGTACGTGTTCGGACATGGGCTCTTCTCGGTAGC	1560
Qy	1561	GGCGGAGTTCCACATCCCGAGCCCTGTGTCCATCCGTCACAGCGCTCATGGTCGCTCGC	1620
Db	1561	GGCGGAGTTCCACATCCCGAGCCCTGTGTCCATCCGTCACAGCGCTCATGGTCGCTCGC	1620
Qy	1621	AGTCTCTTGCTCTTAACAGTGGAGGCGCAGACTTAGGCACAGCAATGCCACCACCAACC	1680
Db	1621	AGTCTCTTGCTCTTAACAGTGGAGGCGCAGACTTAGGCACAGCAATGCCACCACCAACC	1680
Qy	1681	AGTGTCCGCGACAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCGGAGATTGG	1740
Db	1681	AGTGTCCGCGACAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCGGAGATTGG	1740
Qy	1741	GCTCGCACCTGGACGAGATGGAAGACTTAAGCGACGGCAGAGAAGATGCAGGCAGCT	1800
Db	1741	GCTCGCACCTGGACGAGATGGAAGACTTAAGCGACGGCAGAGAAGATGCAGGCAGCT	1800
Qy	1801	GAGTTGTTGATTCTGATAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTAAACGGTGG	1860
Db	1801	GAGTTGTTGATTCTGATAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTAAACGGTGG	1860
Qy	1861	AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGGCCACACAGACATAATAGCT	1920
Db	1861	AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGGCCACACAGACATAATAGCT	1920
Qy	1921	GACAGACTAACAGACTGTTCCCTTTCCATGGTCTTTTCTGAGTCAACGTCGTCGACGAA	1980
Db	1921	GACAGACTAACAGACTGTTCCCTTTCCATGGTCTTTTCTGAGTCAACGTCGTCGACGAA	1980
Qy	1981	TTCAAGCAATCATGGATGCAATGAAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAG	2040
Db	1981	TTCAAGCAATCATGGATGCAATGAAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAG	2040
Qy	2041	CAGTCTTCGTTTCGCCAGCGCTAGCGAAACCAACCGTACCGGGGGAAAGTCCCGGCCACA	2100
Db	2041	CAGTCTTCGTTTCGCCAGCGCTAGCGAAACCAACCGTACCGGGGGAAAGTCCCGGCCACA	2100
Qy	2101	CTGTGCTGCAATTTGTTAGCTCCTCGCACAGGCGCCAGACAGACGTCGACTGATCA	2160
Db	2101	CTGTGCTGCAATTTGTTAGCTCCTCGCACAGGCGCCAGACAGACGTCGACTGATCA	2160
Qy	2161	ACACCAACGCGAGTTGGACCTCAATAGCACGGCCCTGAACTGCAATGATAGCCTCAACA	2220

|||||
2161 ACACCAACGCGAGTTGGCACTCAATAGACACGCGCCCTGAACCTGCAATGATAGCTCAACA 2220
|||||
2221 CCGGCTGGTTGGCAGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCTTGAGA 2280
|||||
2221 CCGGCTGGTTGGCAGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCTTGAGA 2280
|||||
2281 GGCTAGCAGCTGCCGACCCCTTACCGATTTTGAACAGGCTGGGCCCCCTATCAGTTATG 2340
|||||
2281 GGCTAGCAGCTGCCGACCCCTTACCGATTTTGAACAGGCTGGGCCCCCTATCAGTTATG 2340
|||||
2341 CCAACGGAAGCGGCCCGACCGCTACTCTGCTGGCACTACCCCCCAAAACCTTGGC 2400
|||||
2341 CCAACGGAAGCGGCCCGACCGCTACTCTGCTGGCACTACCCCCCAAAACCTTGGC 2400
|||||
2401 GTATTGTGCCCGCAAGAGTGTGTGCTCCGCTATATTGCTTCACTCCACGCCCGCTGG 2460
|||||
2401 GTATTGTGCCCGCAAGAGTGTGTGCTCCGCTATATTGCTTCACTCCACGCCCGCTGG 2460
|||||
2461 TGGTGGAAACGACCGACAGGTGGGCGGCCCACTACAGCTGGGTGAAAATGATACGG 2520
|||||
2461 TGGTGGAAACGACCGACAGGTGGGCGGCCCACTACAGCTGGGTGAAAATGATACGG 2520
|||||
2521 AGCTCTTCCTTAAACAATACAGGCCACCGCTGGGCAATTGGTTGGTTGTATCCCTGGA 2580
|||||
2521 AGCTCTTCCTTAAACAATACAGGCCACCGCTGGGCAATTGGTTGGTTGTATCCCTGGA 2580
|||||
2581 TCAACTCAACTCGATTCACCAAGTGTGGAGCGCTCTCTTGTGTCATCGGAGGGCGG 2640
|||||
2581 TCAACTCAACTCGATTCACCAAGTGTGGAGCGCTCTCTTGTGTCATCGGAGGGCGG 2640
|||||
2641 GCAACAACACCTGCACTGCCCACTGATTGCTTCGCAAGCATCCGACGCCACATACT 2700
|||||
2641 GCAACAACACCTGCACTGCCCACTGATTGCTTCGCAAGCATCCGACGCCACATACT 2700
|||||
2701 CTCGGTGGCTCCGGTCCCTGGATCAACACAGGTGCTGTGCACTACCCGTATAGGC 2760
|||||
2701 CTCGGTGGCTCCGGTCCCTGGATCAACACAGGTGCTGTGCACTACCCGTATAGGC 2760
|||||
2761 TTTGGCATTTCTTGTACCATCACTACACCATATTTAAATCAGGATGACGTGGGAG 2820
|||||
2761 TTTGGCATTTCTTGTACCATCACTACACCATATTTAAATCAGGATGACGTGGGAG 2820
|||||
2821 GGGTGGAAACAGGCTGGAGCTGCTGCAACTGGACGGGGCGCAACCTTGCGATCTGG 2880
|||||
2821 GGGTGGAAACAGGCTGGAGCTGCTGCAACTGGACGGGGCGCAACCTTGCGATCTGG 2880
|||||
2881 AAGATAGGAACAGGCTCCGAGATCGATATGGAGAAACATCAATCAGGATTCCTAGGACCC 2940
|||||
2881 AAGATAGGAACAGGCTCCGAGATCGATATGGAGAAACATCAATCAGGATTCCTAGGACCC 2940
|||||
2941 TGCTCGTTTACAGGGGGGTTTTCTTGTGCAAGATTCCTCAATACCGCAGAGTC 3000
|||||
2941 TGCTCGTTTACAGGGGGGTTTTCTTGTGCAAGATTCCTCAATACCGCAGAGTC 3000
|||||
3001 TAGACTCGTGGTGGACTCTCTCAATTTTCTAGGGGATCTCCCGTGTCTTTGGCCAAA 3060
|||||
3001 TAGACTCGTGGTGGACTCTCTCAATTTTCTAGGGGATCTCCCGTGTCTTTGGCCAAA 3060
|||||
3061 ATTGCGAGTCCCAACCTCCAATCACTCAACACCTCTCTCTCTCAATTTGTCTGGTT 3120
|||||
3061 ATTGCGAGTCCCAACCTCCAATCACTCAACACCTCTCTCTCTCAATTTGTCTGGTT 3120
|||||
3121 ATCGCTGGATGTGTGCGGGGTTTTATCATATTTCTCTCATCTCTGCTATGCTCA 3180
|||||
3121 ATCGCTGGATGTGTGCGGGGTTTTATCATATTTCTCTCATCTCTGCTATGCTCA 3180
|||||
3181 TCTTCTTATTTGTTCTTCTGGATTAACAAGGTATGTTGCGGTTTGTCTCTAAATCCAG 3240
|||||
3181 TCTTCTTATTTGTTCTTCTGGATTAACAAGGTATGTTGCGGTTTGTCTCTAAATCCAG 3240
|||||
3241 GATCAACAAACCAAGTACGGGACCAATGCAAAACCTGCAAGCTCTCTGCTCAAGGCACT 3300
|||||

Db
3241 GATCAACAAACCAAGTACGGGACCATGCAAAACCTGCAAGCTCTCTGCTCAAGGCACT 3300
QY
3301 CTATGTTTCCCTCATGTTGCTGTACAAAACCTACCGATGAAAATTCGACTGTATTCCCA 3360
Db
3301 CTATGTTTCCCTCATGTTGCTGTACAAAACCTACCGATGAAAATTCGACTGTATTCCCA 3360
QY
3361 TCCCATCGTCTGGGCTTTTCGCAAAATACCTATGGAAGTGGGCTCAGTCCGTTCTCTT 3420
Db
3361 TCCCATCGTCTGGGCTTTTCGCAAAATACCTATGGAAGTGGGCTCAGTCCGTTCTCTT 3420
QY
3421 GGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTCTGTTAGGGCTTTCCCACTGTTTGGC 3480
Db
3421 GGCTCAGTTTACTAGTGCATTTGTTTCTAGTGTCTGTTAGGGCTTTCCCACTGTTTGGC 3480
QY
3481 TTTTCACTATATGATGATGTTGTTGTTGGGCGCCAAAGTCTGTACAGCATCGTGAGTCCCT 3540
Db
3481 TTTTCACTATATGATGATGTTGTTGTTGGGCGCCAAAGTCTGTACAGCATCGTGAGTCCCT 3540
QY
3541 TTATACCGCTGTTTACAAATTTTCTTTGCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
Db
3541 TTATACCGCTGTTTACAAATTTTCTTTGCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
QY
3601 AGCAAGTCTAGAAAGCGCGCCCAAGATATCAAGGATCCACTACGGTTAGAGCTCCCTGA 3660
Db
3601 AGCAAGTCTAGAAAGCGCGCCCAAGATATCAAGGATCCACTACGGTTAGAGCTCCCTGA 3660
QY
3661 TCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGCTGCCT 3720
Db
3661 TCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGCTGCCT 3720
QY
3721 TCTTTGACCTCGAAAGGTGCCACTCCACTGCTCTTTCTTAATAAATGAGGAAATTCGA 3780
Db
3721 TCTTTGACCTCGAAAGGTGCCACTCCACTGCTCTTTCTTAATAAATGAGGAAATTCGA 3780
QY
3781 TCGCATTTCTCAGTAGTGTCTATTTCTTCTGTTGTTGGGTTGGGTCAGGACAGCAG 3840
Db
3781 TCGCATTTCTCAGTAGTGTCTATTTCTTCTGTTGTTGGGTTGGGTCAGGACAGCAG 3840
QY
3841 GGGGAGGATTTGGAAAGACAATAGCAGGATGCTGGGAGCTCTTCGCTTCTCTCGCTCAC 3900
Db
3841 GGGGAGGATTTGGAAAGACAATAGCAGGATGCTGGGAGCTCTTCGCTTCTCTCGCTCAC 3900
QY
3901 TGACTCGCTGCTCGCTCGTCTGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGT 3960
Db
3901 TGACTCGCTGCTCGCTCGTCTGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGGT 3960
QY
3961 AATAAGGTTATCCACAGATCAGGGGATTAACCGAGGAAAGACATGTGAGCAAAAGCCCA 4020
Db
3961 AATAAGGTTATCCACAGATCAGGGGATTAACCGAGGAAAGACATGTGAGCAAAAGCCCA 4020
QY
4021 GCAAAAGGCGAGGAAACCGTAAAGCGCGCTTCTGCTGGGCTTTTTCATAGGCTCCGCC 4080
Db
4021 GCAAAAGGCGAGGAAACCGTAAAGCGCGCTTCTGCTGGGCTTTTTCATAGGCTCCGCC 4080
QY
4081 CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAGACT 4140
Db
4081 CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAGACT 4140
QY
4141 ATAAAGATACAGGGGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCTCTGTTCCGACCT 4200
Db
4141 ATAAAGATACAGGGGTTTCCCTCTGGAAGCTCCCTGTCGCTCTCTCTGTTCCGACCT 4200
QY
4201 GCGCTTACCGGATACCTGTCGCCCTTTCTCTCTGGGAAGCGTGGCGCTTTCTCAATG 4260
Db
4201 GCGCTTACCGGATACCTGTCGCCCTTTCTCTCTGGGAAGCGTGGCGCTTTCTCAATG 4260
QY
4261 CTCACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCAAGCTGGGCTGTGTGCA 4320
Db
4261 CTCACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCCTCAAGCTGGGCTGTGTGCA 4320
QY
4321 CGAAACCCCGTTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTCTTGAGTCCAA 4380
Db
4321 CGAAACCCCGTTTACGCGGACCGCTGCGCTTATCCCGTAACTATCGTCTTGAGTCCAA 4380

```
QY 4381 CCGGTAAGACACGACCTTATCGCCACTGCGACGAGCCACTGTAACAGAGTTAGCAGGC 4440
DB 4381 CCGGTAAGACACGACCTTATCGCCACTGCGACGAGCCACTGTAACAGAGTTAGCAGGC 4440
QY 4441 GAGGTATGAGGGCGTGTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAG 4500
DB 4441 GAGGTATGAGGGCGTGTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAG 4500
QY 4501 AAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGG 4560
DB 4501 AAGGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGG 4560
QY 4561 TAGCTCTTGATCCGCAAAACCAACACCGCTGTAGCGGTGGTTTTTTTGGTTGCAAGCA 4620
DB 4561 TAGCTCTTGATCCGCAAAACCAACACCGCTGTAGCGGTGGTTTTTTTGGTTGCAAGCA 4620
QY 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC 4680
DB 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC 4680
QY 4681 TGAAGCTCAGTGAACGAAACCTCACGTTTAAGGATTTTGGTCATGAGATTATCAAAAAAG 4740
DB 4681 TGAAGCTCAGTGAACGAAACCTCACGTTTAAGGATTTTGGTCATGAGATTATCAAAAAAG 4740
QY 4741 GATCTTCACTAGATCCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATATA 4800
DB 4741 GATCTTCACTAGATCCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATATA 4800
QY 4801 TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT 4860
DB 4801 TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT 4860
QY 4861 CTGTCTATTTCGTTTCACTAGTTCGCTGACTCCCGCTCGTGTAGATAACTACGATACG 4920
DB 4861 CTGTCTATTTCGTTTCACTAGTTCGCTGACTCCCGCTCGTGTAGATAACTACGATACG 4920
QY 4921 GAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCAGACCCGCTCACCAGC 4980
DB 4921 GAGGGCTTACCATCTGGCCCGCAGTCTGCAATGATACCGCAGACCCGCTCACCAGC 4980
QY 4981 TCCAGATTATCAGCAATAAACCAGCAGCCGGAAGCGCGAGCAAGTGTCTCTGC 5040
DB 4981 TCCAGATTATCAGCAATAAACCAGCAGCCGGAAGCGCGAGCAAGTGTCTCTGC 5040
QY 5041 AACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAGTTC 5100
DB 5041 AACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAGTTC 5100
QY 5101 GCCAGTTAATGTTGGCAACGTTGTTGCCATTTGCTACAGGCATCGTGGTCAAGCTC 5160
DB 5101 GCCAGTTAATGTTGGCAACGTTGTTGCCATTTGCTACAGGCATCGTGGTCAAGCTC 5160
QY 5161 GTCGTTTGGTATGGCTTTCATTCAGCTCCGCTTCCCAACGATCAAGCGGAGTTACATGATC 5220
DB 5161 GTCGTTTGGTATGGCTTTCATTCAGCTCCGCTTCCCAACGATCAAGCGGAGTTACATGATC 5220
QY 5221 CCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGCTCCGATCGTTGTTCAGAAAGTAA 5280
DB 5221 CCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGCTCCGATCGTTGTTCAGAAAGTAA 5280
QY 5281 GTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATCTTCTACTGTCAT 5340
DB 5281 GTTGGCGCAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATCTTCTACTGTCAT 5340
QY 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCTTCTGAGAAATA 5400
DB 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCTTCTGAGAAATA 5400
QY 5401 GTGTATCGCGCAGCAGGTTGCTCTTTGCCGGCGTCAATACGGGATAATACCGCGCCACA 5460
DB 5401 GTGTATCGCGCAGCAGGTTGCTCTTTGCCGGCGTCAATACGGGATAATACCGCGCCACA 5460
```

```
QY 5461 TAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCTCGGGCGGAAAACTCTCAAG 5520
DB 5461 TAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCTCGGGCGGAAAACTCTCAAG 5520
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGTGCAACCAACTGATCTTC 5580
DB 5521 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGTGCAACCAACTGATCTTC 5580
QY 5581 AGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGAAAAATCGCGC 5640
DB 5581 AGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGAAAAATCGCGC 5640
QY 5641 AAAAAAGGAAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCTTTTCAATA 5700
DB 5641 AAAAAAGGAAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCTTTTCAATA 5700
QY 5701 TTATTGAACATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
DB 5701 TTATTGAACATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
QY 5761 GAAAAATAAACAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCCACGTGACTA 5820
DB 5761 GAAAAATAAACAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCCACGTGACTA 5820
QY 5821 AGAACCATTATTATCATGACATTAACCTATAAATAAGCGGTATCACGAGCCCTTTCG 5880
DB 5821 AGAACCATTATTATCATGACATTAACCTATAAATAAGCGGTATCACGAGCCCTTTCG 5880
QY 5881 TC 5882
DB 5881 TC 5882

RESULT 2
AX154499
LOCUS AX154499 5459 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 4 from Patent WO0138358.
ACCESSION AX154499
VERSION AX154499.1 GI:14536109
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM Selby, M.C., Glazer, E.C. and Houghton, M.C.
REFERENCE 1 Hbv/hcv virus-like particle
AUTHORS Patent: WO 0138358-A 4 31-MAY-2001;
TITLE CHIRON CORPORATION (US)
JOURNAL Location/Qualifiers
FEATURES
source 1..5459
/mol_type="synthetic construct"
/db_xref="taxon:32630"
/translation="MDAMKRGICCVLLLCGAVFSPASQYVRNSTGLVHVTNDCPNS
SIVYEADAALHTPCVPCVREGNARCWVAMTPVATRDGKLPATOLREHIDLVLGS
ATLCSALYVDGLCGSVFLVQGFITSPRRHWTTCGNCNCSYFGHITGHRMAMDMNWN
SPTNSHSPSCPTPCFGRWMLRRFIIFLLILLIIFLLVLLYQGMPLVPLIPG
SPTTGTGPKCTTPAQGNMFPSCCTKPTDGNCTCIPSSWAFKYLWESVRF
SWLSLLVPVQMFVGLSPVTWLSAIWMWYWGSPSLYSIVSPFIPULLPIFFCLWVYI"
CDS
Query Match 78.5%; Score 4617.2; DB 6; Length 5459;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 5241; Conservative 0; Mismatches 218; Indels 423; Gaps 7;
```


QY 1 TCGCGGCTTTCGGTGTATGACGGTGAACAACTCTGACACATGACGCTCCCGGAGACGGTCA 60
DB 1 TCGCGGCTTTCGGTGTATGACGGTGAACAACTCTGACACATGACGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAAGCCGTCAGGCGCGGTGAGCGGGTG 120
DB 61 CAGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAAGCCGTCAGGCGCGGTGAGCGGGTG 120
QY 121 TTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB 121 TTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
QY 181 ACCATATGAAGCTTTTTCAAAAGCCTAGGCTCCAAAAGCCTCCTCACTACTTCG 240
DB 181 ACCATATGAAGCTTTTTCAAAAGCCTAGGCTCCAAAAGCCTCCTCACTACTTCG 240
QY 241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCGATTAATAAAAAAATTAGTCAGCCA 300
DB 241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCGATTAATAAAAAAATTAGTCAGCCA 300
QY 301 TGGGCGGAGATGGCGGAACTGGGCGGGAGGGAATTATTGGCTATTGGCCATTGCA 360
DB 301 TGGGCGGAGATGGCGGAACTGGGCGGGAGGGAATTATTGGCTATTGGCCATTGCA 360
QY 361 AGTTGTATCTATATCAATAATATGATATTTATTTATTTGGCTCAATGACCGCCA 420
DB 361 AGTTGTATCTATATCAATAATATGATATTTATTTGGCTCAATGACCGCCA 420
QY 421 TGTGTACATTGATTATGACTAGTTATTAATAGTAATCAATTAAGGGGTCAATTAGTTCA 480
DB 421 TGTGTACATTGATTATGACTAGTTATTAATAGTAATCAATTAAGGGGTCAATTAGTTCA 480
QY 481 AGCCATATATGAGTTTCGCGTTACATACTTACGTTAAATGGCCGCTGGCTGACCG 540
DB 481 AGCCATATATGAGTTTCGCGTTACATACTTACGTTAAATGGCCGCTGGCTGACCG 540
QY 541 CCCAAGACCCCGCCCATTTGACGTCATTAATGACGATGTTCCCATAGTAGCCCAATA 600
DB 541 CCCAAGACCCCGCCCATTTGACGTCATTAATGACGATGTTCCCATAGTAGCCCAATA 600
QY 601 GGGACTTTCCATTGACGTCATTAAGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA 660
DB 601 GGGACTTTCCATTGACGTCATTAAGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA 660
QY 661 CATCAAGTGTATCATATGCAAGTCCGCCCCCTATTGACGTCATTAAGCGGTAAATGGCCC 720
DB 661 CATCAAGTGTATCATATGCAAGTCCGCCCCCTATTGACGTCATTAAGCGGTAAATGGCCC 720
QY 721 GCGTGCATTATGCCCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
DB 721 GCGTGCATTATGCCCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
QY 781 GTATTAGTCATCGCTAATACATGGTGTATGCGGTTTGGCAGTACACCAATGGCGGTGA 840
DB 781 GTATTAGTCATCGCTAATACATGGTGTATGCGGTTTGGCAGTACACCAATGGCGGTGA 840
QY 841 TAGCGGTTGACTCAGCGGGATTTCCAAAGTCTCCACCCATTGACGTCATTAAGGGAGTTG 900
DB 841 TAGCGGTTGACTCAGCGGGATTTCCAAAGTCTCCACCCATTGACGTCATTAAGGGAGTTG 900
QY 901 TTTTGGCACCAAAATCAAGGGACTTTCCAAAGTCTCCAAATAGCCCGCCCGTTGACG 960
DB 901 TTTTGGCACCAAAATCAAGGGACTTTCCAAAGTCTCCAAATAGCCCGCCCGTTGACG 960
QY 961 CAAATGGCGGTAGCGGTGACGGTGGAGGCTTATATAGCAGAGCTCGTTTGTAGTGAAC 1020
DB 961 CAAATGGCGGTAGCGGTGACGGTGGAGGCTTATATAGCAGAGCTCGTTTGTAGTGAAC 1020
QY 1021 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTGAATCCATAGAGACACCGGGAC 1080
DB 1021 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTGAATCCATAGAGACACCGGGAC 1080
QY 1081 CGATCCAGCTCCGCGCGGGAAACGGTGTGATTGGAAACGGGATTTCCCGTGCAGAGT 1140

DB 1081 CGATCCAGCTCCGCGCGGGAAACGGTGTGATTGGAAACGGGATTTCCCGTGCAGAGT 1140
QY 1141 GACGTAAGTACCGCTATAGACTCTATAGCAGACACCCCTTTGGCTCTTATGCAATGCTATA 1200
DB 1141 GACGTAAGTACCGCTATAGACTCTATAGCAGACACCCCTTTGGCTCTTATGCAATGCTATA 1200
QY 1201 CTGTTTTTGGCTTTGGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGTATAGCT 1260
DB 1201 CTGTTTTTGGCTTTGGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGTATAGCT 1260
QY 1261 TAGCCTATAGGTGTTGGGTTATTGAACATTTATGAACACTCCCTTATGCTATAGGTGATGACTTT 1320
DB 1261 TAGCCTATAGGTGTTGGGTTATTGAACATTTATGAACACTCCCTTATGCTATAGGTGATGACTTT 1320
QY 1321 TCCATTACTAATCCATAACATGCTCTTTGCGACAACTATCTCTATTTGGCTATATGCCAA 1380
DB 1321 TCCATTACTAATCCATAACATGCTCTTTGCGACAACTATCTCTATTTGGCTATATGCCAA 1380
QY 1381 TACTCTGCTCTTACAGAGCTGACACGGAATCTGTATTTTTTACAGGATGGGCTCCATTAT 1440
DB 1381 TACTCTGCTCTTACAGAGCTGACACGGAATCTGTATTTTTTACAGGATGGGCTCCATTAT 1440
QY 1441 TATTTACAAATTCACATATACAAACGCGCTCCCGCTGCGCAGTATTTTATTAACA 1500
DB 1441 TATTTACAAATTCACATATACAAACGCGCTCCCGCTGCGCAGTATTTTATTAACA 1500
QY 1501 TAGCGTGGGATCTCCGACATCTCGGTAGTGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
DB 1501 TAGCGTGGGATCTCCGACATCTCGGTAGTGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
QY 1561 GCGGAGCTTCCACATCCGAGCCTCGTCCCATCCGTCAGCGGCTCATGGTCCGCTCGGC 1620
DB 1561 GCGGAGCTTCCACATCCGAGCCTCGTCCCATCCGTCAGCGGCTCATGGTCCGCTCGGC 1620
QY 1621 AGTCTCTGCTCTTAAACAGTGGAGCCAGATTTAGGCAAGCAATGCCACACACACC 1680
DB 1621 AGTCTCTGCTCTTAAACAGTGGAGCCAGATTTAGGCAAGCAATGCCACACACACC 1680
QY 1681 AGTGTGCCGACAGCGCGGTGGGTATGTTCTGAAATAGCTCGGAGATTGG 1740
DB 1681 AGTGTGCCGACAGCGCGGTGGGTATGTTCTGAAATAGCTCGGAGATTGG 1740
QY 1741 GCTCGCACCTGACGAGATGGAAGTCTTAAGCAGCGCAGAAAGATGACGAGCAGCT 1800
DB 1741 GCTCGCACCTGACGAGATGGAAGTCTTAAGCAGCGCAGAAAGATGACGAGCAGCT 1800
QY 1801 GAGTTGTTGTTATTTGATTAAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTTAACCGTGG 1860
DB 1801 GAGTTGTTGTTATTTGATTAAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTTAACCGTGG 1860
QY 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCCACACAGACATAATAGCT 1920
DB 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCCACACAGACATAATAGCT 1920
QY 1921 GACAGACTAACAGACTGTTCTTCTTCCATGGGTCTTTTCTGCACTACCCGTCGCGAGAA 1980
DB 1921 GACAGACTAACAGACTGTTCTTCTTCCATGGGTCTTTTCTGCACTACCCGTCGCGAGAA 1980
QY 1981 TTTCAAGCAATCATGGATGCAATGAAGAGGGCTCTGCTGTGTGCTGTGTGTGGAG 2040
DB 1981 TTTCAAGCAATCATGGATGCAATGAAGAGGGCTCTGCTGTGTGCTGTGTGTGGAG 2040
QY 2041 CAGTCTTCTGTTTGGCCAGGCTTAGGAAACCCACGTCACCGGGGAGTAGCCGCCACA 2100
DB 2041 CAGTCTTCTGTTTGGCCAGGCTTAGGAAACCCACGTCACCGGGGAGTAGCCGCCACA 2100
QY 2101 CTGTGCTGGAATTTGTTAGCTCTCTCGCACAGCGCCCAAGCAGAGACGTCAGCTGATCA 2160
DB 2069 -----CCAGGTGGCA 2079
QY 2161 ACACCAACGGCAGTTGGCACTCTCAATAGCACGCGCCCTGAACTGCAATGATGAGCTCAACA 2220
DB 2161 ACACCAACGGCAGTTGGCACTCTCAATAGCACGCGCCCTGAACTGCAATGATGAGCTCAACA 2220

Db 2080 ACAGCACCGGCTGTACACGTGA-----CCAAAGACTGCCCAACA 2121
QY 2221 CCGGCTGGTGGCAGGGCTTTCTATCAACAAGATTCACTCTCAGGCTGCTGAGA 2280
Db 2122 GCAGCATCGTGTACGAGGCGCGCGCATCTGACACACCCCGGCTCGTGCC----- 2176
QY 2281 GGCTAGCCAGCTGCCGACCCCTTACCGATTTTGGACCAAGGCTGGGGCCCTATCAGTTATG 2340
Db 2177 -----CTGCTGGCGGAGGCAACGCCAGCCGCTGCTGGGTGGCAATGACCCCA 2226
QY 2341 CCAACGGAAGCGGCCCGACAGCGCCCTTACTGCTGGCACTACCCCCCAAAACCTTTGG 2400
Db 2227 CCGTGGCCACCGCGAGCGCAAGCTGCCCGCACCCAGCTGCGCGC----- 2273
QY 2401 GTATTGTCCCGCGAAGAGTGTGTGGTTCGGGTATATTGCTTCACTCCAGCCCCGTGG 2460
Db 2274 -----CACATCGACCTGC 2286
QY 2461 TGGTGGGAACGACGACAGGTGGGCGCGCCACCTACAGCTGGGTGGAATGATACGG 2520
Db 2287 TGGTGGGAGCGCCACCTGTGAGCGCC-----TGT 2319
QY 2521 ACCTCTTCTGCTTAACAATACAGGCGCACCGCTGGGCAATTGGTTGGTTGTACCTGGA 2580
Db 2320 ACGTGGCGACCTGTGGGCGAGCTGTTCTGGTGGGCGCAGCTGTT----- 2366
QY 2581 TGAATCAATGAATTCAACAAAGTGTGGAGCGCTCTTGTGTATCGGAGGGGCG 2640
Db 2367 -----ACCT 2370
QY 2641 GCAACAACCCCTGACTGCCACACTGATTTCTTCGCAAGCATCGGACGCCACATCT 2700
Db 2371 TCAGCCCCCGCCCACTGACACCCAGCGGTGCAACTGCAAGTCTACCCCGGCCACA 2430
QY 2701 CTGGTGGCGCTCCGCTCCCTGATCACACCCAGGTGCTGTGCGACTACCCCGTATAGGC 2760
Db 2431 TCACCGGCCACCGCATGGCTGGGACATGATGAT----- 2464
QY 2761 TTTGGCATTAATCTTTGATCAATCAACTACACCAATATTAAATCAGGATGTAGTGGGAG 2820
Db 2465 ----- 2464
QY 2821 GGCTCGAACACAGGCTGGAAGCTGCCTGCACTGGACGGGCGCAACGTTGCGATCTGG 2880
Db 2465 ----- 2464
QY 2881 AAGATAGGACAGGTCGAGATCGATATGAGAAACATCACTCAGGATTCCTAGGACCCC 2940
Db 2465 -----GAACCTGGAGCCCAACCATGGAGAACATCACTCAGGATTCCTAGGACCCC 2517
QY 2941 TGCTCGTTTACAGGGGGGTTTTCTTTGTTGAAGAATCTCTCAAAATCCGCAAGTCC 3000
Db 2518 TGCTCGTTTACAGGGGGGTTTTCTTTGTTGAAGAATCTCTCAAAATCCGCAAGTCC 2577
QY 3001 TAGACTCGTGGTGAATCTCTCAATTTCTAGGGGATCTCCGCTGTGCTTTGGCCAAA 3060
Db 2578 TAGACTCGTGGTGAATCTCTCAATTTCTAGGGGATCTCCGCTGTGCTTTGGCCAAA 2637
QY 3061 ATTTCGAGTCCCCAACTCCAAATCACTACCAACCTCCTGCTCCAAATTTGCTGGTT 3120
Db 2638 ATTTCGAGTCCCCAACTCCAAATCACTACCAACCTCCTGCTCCAAATTTGCTGGTT 2697
QY 3121 ATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTCTTTCATCTGCTGCTATGCTCA 3180
Db 2698 ATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTCTTTCATCTGCTGCTATGCTCA 2757
QY 3181 TCTTCTTATTTGTTCTCTGGATTATCAAGTATGTTGCCGTTTCTCTCTAAATTTCCAG 3240
Db 2758 TCTTCTTATTTGTTCTCTGGATTATCAAGTATGTTGCCGTTTCTCTCTAAATTTCCAG 2817
QY 3241 GATCAACAACACAGTACGGGACCATGCAAAACCTGCAGACTCTCTGCTCAAGGCAACT 3300
Db 2818 GATCAACAACACAGTACGGGACCATGCAAAACCTGCAGACTCTCTGCTCAAGGCAACT 2877

QY 3301 CTATGTTTCCCTCATGTTGCTGTACAAACCTACCGATGGAAATTCACCTGTATTCCCA 3360
Db 2878 CTATGTTTCCCTCATGTTGCTGTACAAACCTACCGATGGAAATTCACCTGTATTCCCA 2937
QY 3361 TCCCATCGTCTCGGGCTTTTCGCAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT 3420
Db 2938 TCCCATCGTCTCGGGCTTTTCGCAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT 2997
QY 3421 GGCTCAGTTTATCTAGTGCATTTGTTCACTGTTTCTAGGGTTCCTCCCACTGTTTGGC 3480
Db 2998 GGCTCAGTTTATCTAGTGCATTTGTTCACTGTTTCTAGGGTTCCTCCCACTGTTTGGC 3057
QY 3481 TTTTACGATATATGGATGATGTTATTTGGGGGCCAAGTCTGTACAGCATCGTGAGTCCCT 3540
Db 3058 TTTTACGATATATGGATGATGTTATTTGGGGGCCAAGTCTGTACAGCATCGTGAGTCCCT 3117
QY 3541 TTTATACCGCTGTATCCAAATTTCTTTGTTCTCTGGGTATACATTTTAAAGAAATTCAGACTCG 3600
Db 3118 TTTATACCGCTGTATCCAAATTTCTTTGTTCTCTGGGTATACATTTTAAAGAAATTCAGACTCG 3177
QY 3601 AGCAAGTCTAGAAAAGCGCGCAAGATATCAAGGATCCACTACGCTTTAGAGTCTCGCTGA 3660
Db 3178 AGCAAGTCTAGAAAAGCGCGCAAGATATCAAGGATCCACTACGCTTTAGAGTCTCGCTGA 3237
QY 3661 TCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGCTGCCCT 3720
Db 3238 TCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGCTGCCCT 3297
QY 3721 TCCTTGACCTCGAAGTGCCTCCACTGCTCTTCTCTTAATAAGGAAATTTGCA 3780
Db 3298 TCCTTGACCTCGAAGTGCCTCCACTGCTCTTCTTAATAAGGAAATTTGCA 3357
QY 3781 TCGCATTTCTCAGTAGTGTCTATTCTATTCTGCGGGGTGGGTGGGAGGACAGCAAG 3840
Db 3358 TCGCATTTCTCAGTAGTGTCTATTCTATTCTGCGGGGTGGGTGGGAGGACAGCAAG 3417
QY 3841 GGGAGGATTTGGAGACAAATAGCAGGCATCTGGGAGCTCTTCCGCTTCCCTCGCTCAC 3900
Db 3418 GGGAGGATTTGGAGACAAATAGCAGGCATCTGGGAGCTCTTCCGCTTCCCTCGCTCAC 3477
QY 3901 TGACTCGCTCGCTCGCTCGTTCCGCTGGCGAGCGGTATCAGCTCACTCAAAAGCGGT 3960
Db 3478 TGACTCGCTCGCTCGCTCGTTCCGCTGGCGAGCGGTATCAGCTCACTCAAAAGCGGT 3537
QY 3961 AATACGGTTTATCCACAGAATCAGGGATTAACGAGGAAAGAAACATGTAGCAAAAGGCCA 4020
Db 3538 AATACGGTTTATCCACAGAATCAGGGATTAACGAGGAAAGAAACATGTAGCAAAAGGCCA 3597
QY 4021 GCAAAAGGCGAGGACCGTAAAGGCGCGTTGCTGGCGTTTTCATAGGCTCCGCC 4080
Db 3598 GCAAAAGGCGAGGACCGTAAAGGCGCGTTGCTGGCGTTTTCATAGGCTCCGCC 3657
QY 4081 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGGATGGGAAACCCGACAGGACT 4140
Db 3658 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGGATGGGAAACCCGACAGGACT 3717
QY 4141 ATAAAGATACAGGGGTTTTCCCTCTGGAAGTCCCTCTGTCGCTCTCTGTTCCGACCCCT 4200
Db 3718 ATAAAGATACAGGGGTTTTCCCTCTGGAAGTCCCTCTGTCGCTCTCTGTTCCGACCCCT 3777
QY 4201 GCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGGCTGGCTTCTCAATG 4260
Db 3778 GCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGGCTGGCTTCTCAATG 3837
QY 4261 CTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCTGGGCTGTGTGCA 4320
Db 3838 CTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCAAGCTGGGCTGTGTGCA 3897
QY 4321 CGAAACCCCGCTTACGCCGAGCGCTGCGCTTATTCGGGTAACTATCTGCTTGAAGTCCAA 4380
Db 3898 CGAAACCCCGCTTACGCCGAGCGCTGCGCTTATTCGGGTAACTATCTGCTTGAAGTCCAA 3957

```
QY 4381 CCCGTAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGATTACGAGC 4440
Db 3958 CCCGGTAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGGTAACAGGATTACGAGC 4017
QY 4441 GAGGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGCCTTAACATACGGCTACACTAG 4500
Db 4018 GAGGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGCCTTAACATACGGCTACACTAG 4077
QY 4501 AAGGACAGATATTTGGTATCTGCCCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGG 4560
Db 4078 AAGGACAGATATTTGGTATCTGCCCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGG 4137
QY 4561 TAGCTCTTGATCGCGCAACAAACACCGCTGTAGCGGTGGTTTTTTTTTTGTTTGAAGCA 4620
Db 4138 TAGCTCTTGATCGCGCAACAAACACCGCTGTAGCGGTGGTTTTTTTTTTGTTTGAAGCA 4197
QY 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACGGGGTC 4680
Db 4198 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACGGGGTC 4257
QY 4681 TGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGAGATTATCAAAAAG 4740
Db 4258 TGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGAGATTATCAAAAAG 4317
QY 4741 GATCTTTCACTAGATCCCTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATA 4800
Db 4318 GATCTTTCACTAGATCCCTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATA 4377
QY 4801 TGAGTAAACTTGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGAT 4860
Db 4378 TGAGTAAACTTGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGAT 4437
QY 4861 CTGCTCTATTCTGTTTCAATAGTCTGCTGACCTCCCGCTGCTGATCACTACGATACG 4920
Db 4438 CTGCTCTATTCTGTTTCAATAGTCTGCTGACCTCCCGCTGCTGATCACTACGATACG 4497
QY 4921 GGAGGCTTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCCAGCTCACCAGC 4980
Db 4498 GGAGGCTTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCCAGCTCACCAGC 4557
QY 4981 TCCAGATTTATCAGCAATTAACAGCGAGCGGGAAGCGCGAGCAAGTGGTCTCTGC 5040
Db 4558 TCCAGATTTATCAGCAATTAACAGCGAGCGGGAAGCGCGAGCAAGTGGTCTCTGC 4617
QY 5041 AACTTTATCGCTCCATCAGCTCTATTATCTGTCGCGGAAGCTAGAGTAGTCTC 5100
Db 4618 AACTTTATCGCTCCATCAGCTCTATTATCTGTCGCGGAAGCTAGAGTAGTCTC 4677
QY 5101 GCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCACTCGTGGTGTCAAGCTC 5160
Db 4678 GCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCACTCGTGGTGTCAAGCTC 4737
QY 5161 GTCGTTTGGTATGGCTTCAATCAGCTCCGGTTCCCAACGATCAAGCGGAGTTACATGATC 5220
Db 4738 GTCGTTTGGTATGGCTTCAATCAGCTCCGGTTCCCAACGATCAAGCGGAGTTACATGATC 4797
QY 5221 CCCCATGTTGTCAAAAAGCGGTAGCTCTTCGCTCCGCTCCGATCGTTGTCAAGATGA 5280
Db 4798 CCCCATGTTGTCAAAAAGCGGTAGCTCTTCGCTCCGCTCCGATCGTTGTCAAGATGA 4857
QY 5281 GTTGGCGCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATCTCTTACTGTCTAT 5340
Db 4858 GTTGGCGCGAGTGTATCACTCATGTTTATGGCAGCACTGCATAATCTCTTACTGTCTAT 4917
QY 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 5400
Db 4918 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 4977
QY 5401 GTCTATGCGCGCAGCCAGTGTCTCTGCGCGGTCAATACCGGATTAATACCGCGCCACA 5460
Db 4978 GTGTATGCGCGCAGCCAGTGTCTCTGCGCGGTCAATACCGGATTAATACCGCGCCACA 5037
QY 5461 TAGCAGAACTTTAAAAGTGTCTCATCATTTGAAAAACGTTCTTTCGGGGCGAAAACTCTCAAG 5520
```

```
Db 5038 TAGCAGAACTTTAAAAGTGTCTCATCATTTGAAAAACGTTCTTTCGGGGCGAAAACTCTCAAG 5097
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTCGTCACCACTGATCTTC 5580
Db 5098 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTCGTCACCACTGATCTTC 5157
QY 5581 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAAAGGAAAGGCAAAATGCGCG 5640
Db 5158 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAAAGGAAAGGCAAAATGCGCG 5217
QY 5641 AAAAAAGGAAATAAGGCGCACACGGAATGTTGAATATCTCATACTCTTCTCTTTCAATA 5700
Db 5218 AAAAAAGGAAATAAGGCGCACACGGAATGTTGAATATCTCATACTCTTCTCTTTCAATA 5277
QY 5701 TTATTGAACGATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
Db 5278 TTATTGAACGATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTA 5337
QY 5761 GAAAAATAAAACAAATAGGCGTTCGCGCACATTTCCCGCAAAAGTGCCACCTCAGCGTCTA 5820
Db 5338 GAAAAATAAAACAAATAGGCGTTCGCGCACATTTCCCGCAAAAGTGCCACCTCAGCGTCTA 5397
QY 5821 AGAAACCATTTATTCATGACATTAACCTTATAAAATAGCGGTATCAGGAGCCCTTTTCG 5880
Db 5398 AGAAACCATTTATTCATGACATTAACCTTATAAAATAGCGGTATCAGGAGCCCTTTTCG 5457
QY 5881 TC 5882
Db 5458 TC 5459

RESULT 3
AX154497 LOCUS
DEFINITION Sequence 2 from Patent WO0138358.
ACCESSION AX154497
VERSION AX154497.1 GI:14536107
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Selby, M. C., Glazer, E. C. and Houghton, M. C.
TITLE Hbv/hcv virus-like particle
JOURNAL Patent: WO 0138358-A 2 31-MAY-2001;
CHIRON CORPORATION (US)
FEATURES
    source
        1. 5128
            Location/Qualifiers
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="plasmid pCMVII-ps2-Sag"
                1988. .2833
                /note="unnamed protein product"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC42600.1"
                /db_xref="GI:14536108"
CDS
    1
        1988. .2833
            /note="unnamed protein product"
            /transl_table=11
            /protein_id="CAC42600.1"
            /db_xref="GI:14536108"
ORIGIN
    Query Match 71.2%; Score 4190; DB 6; Length 5128;
    Best Local Similarity 85.7%; Pred. No. 0;
    Matches 5038; Conservative 0; Mismatches 90; Indels 754; Gaps 4;
    QY 1 TCAGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACAGCTCCCGGAGACGGTCA 60
    |||||
```

Db 1 TCGCGCGTTTCGGTGATGACGGTGAAAACTCTGACACATGTCAGCTCCCGGAGACGGTCA 60
Qy 61 CAGCTTGTCTGTAAAGCGGATGCGGGAGACAGAACAAGCCCGTCAAGGCGCGTCAAGCGGGTG 120
Db 61 CAGCTTGTCTGTAAAGCGGATGCGGGAGACAGAACAAGCCCGTCAAGGCGCGTCAAGCGGGTG 120
Qy 121 TTGGCGGGTGTGGGGTCTGCTTAACTATGCGGCATCAGAGCAGATTGTAAGAGATGC 180
Db 121 TTGGCGGGTGTGGGGTCTGCTTAACTATGCGGCATCAGAGCAGATTGTAAGAGATGC 180
Qy 181 ACCATATGAAGCTTTTTCAAAAAGCCCTAGGCCCTCCAAAAAGCCCTCTCACTACTTCTGG 240
Db 181 ACCATATGAAGCTTTTTCAAAAAGCCCTAGGCCCTCCAAAAAGCCCTCTCACTACTTCTGG 240
Qy 241 AATAGCTCAGAGCGCAGCGCGCTCGGCCCTCTGTCATAAATAAAAAAATTAAGTCAAGCCA 300
Db 241 AATAGCTCAGAGCGCAGCGCGCTCGGCCCTCTGTCATAAATAAAAAAATTAAGTCAAGCCA 300
Qy 301 TGGGCGGGAATGGGCGGAACCTGGGCGGGAGGGAATTTTGGCTATTGGCCATTGGCAT 360
Db 301 TGGGCGGGAATGGGCGGAACCTGGGCGGGAGGGAATTTTGGCTATTGGCCATTGGCAT 360
Qy 361 ACCTGTATCTATATCATATATGTACATTTATTTGGCTCATGTCCAAATATGACGCCCA 420
Db 361 ACCTGTATCTATATCATATATGTACATTTATTTGGCTCATGTCCAAATATGACGCCCA 420
Qy 421 TGTGACATTGATTATGACTAGTTAATAAGTAATCAATTAAGGGGTCAATTAGTTTCAT 480
Db 421 TGTGACATTGATTATGACTAGTTAATAAGTAATCAATTAAGGGGTCAATTAGTTTCAT 480
Qy 481 AGCCCATATGAGTTCGCGGTACATAAATGAGCTATGTTCCCATAGTAAGCCCAATA 600
Db 481 AGCCCATATGAGTTCGCGGTACATAAATGAGCTATGTTCCCATAGTAAGCCCAATA 600
Qy 541 CCCAACGACCCCGCCCATTTGACGTCAATAATGAGCTATGTTCCCATAGTAAGCCCAATA 600
Db 541 CCCAACGACCCCGCCCATTTGACGTCAATAATGAGCTATGTTCCCATAGTAAGCCCAATA 600
Qy 601 GGGACTTTCATTTGACGTCAATGGTGGAGTATTTACGGTAAACTGCGGTAAGTGGCAGTA 660
Db 601 GGGACTTTCATTTGACGTCAATGGTGGAGTATTTACGGTAAACTGCGGTAAGTGGCAGTA 660
Qy 661 CATCAAGTGTATCATATGCGCAAGTCCGCGCCCTATTGACGTCAATGACGGTAATGGCCC 720
Db 661 CATCAAGTGTATCATATGCGCAAGTCCGCGCCCTATTGACGTCAATGACGGTAATGGCCC 720
Qy 721 GCCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
Db 721 GCCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
Qy 781 GTATTAGTCAATGCGGTATTACATGGTGTATGCGGTTTGGCAGTACACCAATGGCGGTGA 840
Db 781 GTATTAGTCAATGCGGTATTACATGGTGTATGCGGTTTGGCAGTACACCAATGGCGGTGA 840
Qy 841 TAGCGGTTTACACTCAGGGGATTTCCAAAGTCTCCACCCATTGACGCTCAATGGGAGTTTG 900
Db 841 TAGCGGTTTACACTCAGGGGATTTCCAAAGTCTCCACCCATTGACGCTCAATGGGAGTTTG 900
Qy 901 TTTTGGCACCACCAATCAACGGGACTTTCCAAAATGTGCTAATAACCCCGCCCGCTTGACG 960
Db 901 TTTTGGCACCACCAATCAACGGGACTTTCCAAAATGTGCTAATAACCCCGCCCGCTTGACG 960
Qy 961 CAAATGGCGGTAGCGGTGTACGCGTGGAGGTCTATATAAGCAGAGCTCGTTTGTAGTAAC 1020
Db 961 CAAATGGCGGTAGCGGTGTACGCGTGGAGGTCTATATAAGCAGAGCTCGTTTGTAGTAAC 1020
Qy 1021 CGTCAGATCCGCTGGAGACCCCATCCACGCTGTTTGGACCTCCATAGAAGACACCGGGAC 1080
Db 1021 CGTCAGATCCGCTGGAGACCCCATCCACGCTGTTTGGACCTCCATAGAAGACACCGGGAC 1080
Qy 1081 CGATCAGACCTCCGCGCCCGGGAACCGGTGATTTGGAACCGGGATTTCCCGCTGCGCAAGAT 1140
Db 1081 CGATCAGACCTCCGCGCCCGGGAACCGGTGATTTGGAACCGGGATTTCCCGCTGCGCAAGAT 1140

Qy 1141 GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGATGCTATA 1200
Db 1141 GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGATGCTATA 1200
Qy 1201 CTGTTTTTGGCTTTGGGGCCTATACACCCCGCTCCCTTATGCTATAGCTGATGCTATAGCT 1260
Db 1201 CTGTTTTTGGCTTTGGGGCCTATACACCCCGCTCCCTTATGCTATAGCTGATGCTATAGCT 1260
Qy 1261 TAGCCTATAGGTGTGGGTTATTGACCATTTATTGACCACTCCCCTATTGCTGACGACTTT 1320
Db 1261 TAGCCTATAGGTGTGGGTTATTGACCATTTATTGACCACTCCCCTATTGCTGACGACTTT 1320
Qy 1321 TCCATTACTAATCCATAAATGCTCTTTGCCACAACTATCTCTATTGGCTATATGCCAA 1380
Db 1321 TCCATTACTAATCCATAAATGCTCTTTGCCACAACTATCTCTATTGGCTATATGCCAA 1380
Qy 1381 TACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTTTACAGGATGGGTCATTTAT 1440
Db 1381 TACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTTTACAGGATGGGTCATTTAT 1440
Qy 1441 TATTTTCAAAATTCACATATACAAACGCGCTCCCGCTGCGGCAGTTTATTATTAACA 1500
Db 1441 TATTTTCAAAATTCACATATACAAACGCGCTCCCGCTGCGGCAGTTTATTATTAACA 1500
Qy 1501 TAGCGTGGGATCTCCGACATCTCGGTAAGTTCGGGATGGGCTCTTCTCCGGTAGC 1560
Db 1501 TAGCGTGGGATCTCCGACATCTCGGTAAGTTCGGGATGGGCTCTTCTCCGGTAGC 1560
Qy 1561 GGGGAGCTTTCACATCCGAGCCCTGCTCCATCCGTCAGCGGCTCATGCTCGCTCGGC 1620
Db 1561 GGGGAGCTTTCACATCCGAGCCCTGCTCCATCCGTCAGCGGCTCATGCTCGCTCGGC 1620
Qy 1621 AGCTCCTTCTCTTAAACAGTGGAGCCAGACTTAGGCAACAGCAATGCCACCAACC 1680
Db 1621 AGCTCCTTCTCTTAAACAGTGGAGCCAGACTTAGGCAACAGCAATGCCACCAACC 1680
Qy 1681 AGTGTCCCCACAAAGCCCTGGGTAAGTGTCTGMAAATGAGCTCGGAGATTGG 1740
Db 1681 AGTGTCCCCACAAAGCCCTGGGTAAGTGTCTGMAAATGAGCTCGGAGATTGG 1740
Qy 1741 GCTCGCACCTGACGAGATGGAAGACTTAAAGCAGCGGCAGAAAGATGACGAGCAGCT 1800
Db 1741 GCTCGCACCTGACGAGATGGAAGACTTAAAGCAGCGGCAGAAAGATGACGAGCAGCT 1800
Qy 1801 GAGTTTGTGTATTTCTGATAAGAGTCAAGGTAACCTCCCGTTGCGGTGCTGTTAACGGTGG 1860
Db 1801 GAGTTTGTGTATTTCTGATAAGAGTCAAGGTAACCTCCCGTTGCGGTGCTGTTAACGGTGG 1860
Qy 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACAGACATAATAGCT 1920
Db 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACAGACATAATAGCT 1920
Qy 1921 GACAGACTAACAGACTGTTTCTTCCATGGGTCTTTTCTGCACTACCGCTCGTGCAGAA 1980
Db 1921 GACAGACTAACAGACTGTTTCTTCCATGGGTCTTTTCTGCACTACCGCTCGTGCAGCTA 1980
Qy 1981 TTCAAGCAATTCATGGATGCAATGAAGAGGGGCTGCTGCTGTGTGCTGCTGTGTGGAG 2040
Db 1981 AGAATTCATGCA-----GTGGAA 1998
Qy 2041 CAGTCTTCTGTTCCGCCAGGCTAGCGAAACCCACGTCACCGGGGAGTGCAGGACACA 2100
Db 1999 CTCCTACTGCTTCCACAAACTCTGAGGATCCAGAGTCAGGGGCTCTGTATCTTCTGTC 2058
Qy 2101 CTGTGTCTGGATTGTGTAGCCTCTCCGACACAGGCGCCCAAGCAGAACTCCAGCTATCA 2160
Db 2059 TGGTGGCTC----- 2067
Qy 2161 ACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACTGCAATGATAGCTCAACA 2220
Db 2068 -----CAGTTTCAGGAACAGTAACCCCTGCTCCGAATATTGCT----- 2105

Db 1627 CCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAC 3686
Qy 4441 GAGGTATGATAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCCTAACTACGGCTACACTAG 4500
Db 3687 GAGGTATGATAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCCTAACTACGGCTACACTAG 3746
Qy 4501 AAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGG 4560
Db 3747 AAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTTGG 3806
Qy 4561 TAGCTCTTGATCCGCAACAAACACACCGCTGCTAGCGGTGGTTTTTTTGGTTTGAAGCA 4620
Db 3807 TAGCTCTTGATCCGCAACAAACACACCGCTGCTAGCGGTGGTTTTTTTGGTTTGAAGCA 3866
Qy 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACCGGGTC 4680
Db 3867 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACCGGGTC 3926
Qy 4681 TGACGCTCAGTGAAAGAAACCTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAG 4740
Db 3927 TGACGCTCAGTGAAAGAAACCTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAG 3986
Qy 4741 GATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 4800
Db 3987 GATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 4046
Qy 4801 TGAGTAAACTTGTCTGACAGTTTAAATGCTTAAATCAGTGAGCCACCTATCTCAGCGAT 4860
Db 4047 TGAGTAAACTTGTCTGACAGTTTAAATGCTTAAATCAGTGAGCCACCTATCTCAGCGAT 4106
Qy 4861 CTGTCTATTTCTGATCCATAGTGTGCTGACTCCCGTGGTGTAGATACTACGATACG 4920
Db 4107 CTGTCTATTTCTGATCCATAGTGTGCTGACTCCCGTGGTGTAGATACTACGATACG 4166
Qy 4921 GGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGAGCCACGCTCACCGC 4980
Db 4167 GGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCAGAGCCACGCTCACCGC 4226
Qy 4981 TCAGATTATACAGCAATAAACAGCCAGCCGGAAGCGCGCAGAGTGGTCTCTGC 5040
Db 4227 TCAGATTATACAGCAATAAACAGCCAGCCGGAAGCGCGCAGAGTGGTCTCTGC 4286
Qy 5041 AACTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAGTTC 5100
Db 4287 AACTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAGTTC 4346
Qy 5101 GCCAGTTAATAGTTTGGCAACCTGTTGCAATTTGCTACAGGCATCGTGGTGTACGCTC 5160
Db 4347 GCCAGTTAATAGTTTGGCAACCTGTTGCAATTTGCTACAGGCATCGTGGTGTACGCTC 4406
Qy 5161 GTCGTTTGGTATGGCTTCATTCAGTCCGGTCCCAACGATCAAGCGAGTTACATGATC 5220
Db 4407 GTCGTTTGGTATGGCTTCATTCAGTCCGGTCCCAACGATCAAGCGAGTTACATGATC 4466
Qy 5221 CCCATGTTGTGCAAAAGCGGTTAGCTCTTCCTGCTCCCATCGTTGTCAGAGTAA 5280
Db 4467 CCCATGTTGTGCAAAAGCGGTTAGCTCTTCCTGCTCCCATCGTTGTCAGAGTAA 4526
Qy 5281 GTTGGCGGAGTGTATCACTCATGTTATGCGCAGACTGCATAATTTCTTACTGTCTCAT 5340
Db 4527 GTTGGCGGAGTGTATCACTCATGTTATGCGCAGACTGCATAATTTCTTACTGTCTCAT 4586
Qy 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 5400
Db 4587 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 4646
Qy 5401 GTGTATGCGCGCAGCCAGTGTCTCTTCCCGCGGTCAATACGGGTAATAACCGGCCACA 5460
Db 4647 GTGTATGCGCGCAGCCAGTGTCTCTTCCCGCGGTCAATACGGGTAATAACCGGCCACA 4706
Qy 5461 TAGCAGAACTTTAAAGTGTCTCATCTATTTGAAAAACGTTCTTCGGGGCGAAAACTCTCAAG 5520
Db 4707 TAGCAGAACTTTAAAGTGTCTCATCTATTTGAAAAACGTTCTTCGGGGCGAAAACTCTCAAG 4766

Qy 5521 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATCTTC 5580
Db 4767 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATCTTC 4826
Qy 5581 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGCAAAATCCCGC 5640
Db 4827 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGCAAAATCCCGC 4886
Qy 5641 AAAAAAGGAATAAAGGCGCACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATA 5700
Db 4887 AAAAAAGGAATAAAGGCGCACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATA 4946
Qy 5701 TTATTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTA 5760
Db 4947 TTATTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGATTTA 5006
Qy 5761 GAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGCTTA 5820
Db 5007 GAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGCTTA 5066
Qy 5821 AGAAACCATTTATTATCATGACATTAACCTATATAAATAGCGGTATCACGAGGCCCTTTCG 5880
Db 5067 AGAAACCATTTATTATCATGACATTAACCTATATAAATAGCGGTATCACGAGGCCCTTTCG 5126
Qy 5881 TC 5882
Db 5127 TC 5128

RESULT 4
AX164584
LOCUS
DEFINITION Sequence 6 from Patent WO0138360.
ACCESSION AX164584
VERSION AX164584.1 GI:14545518
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Coit, D.C., Medina-Selby, A.C., Selby, M.C. and Houghton, M.C.
TITLE Novel hcv non-structural polypeptide
JOURNAL Patent: WO 0138360-A 6 31-MAY-2001;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..6299
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pNS34a"
1990..4050
/code="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC42724.1"
/db_xref="GI:14545519"
/translation="MAPITAYAQOTRGLGCIITSITGRDKNQEVEQIVSTAAQTFL
LATCINGVCTVTHGAGTITIASPKGPVITMYTNDQDLVGPASQGTSLTPTCTGCS
SDLYLTVRHADVIPIVRRDGSRLSPRIISYLVKSGSGGPLLCPAGHAGVGPRAAVC
TGAKAVDFIPIVENLETMRSPVTDNSSPPVPQSFQVHLHAPTHGSGSKTKVPA
YAAQGVKVLNPSVAATLGFAYMSKAGIDPNIETGRTITTTGSPITYSTYKFLA
DGGCGGAYDIIICDECHSDATSIILGIGTLDQAEATAGARLVLAALATAPPGSVTVPH
PNIIEVALSTGEIPFYKAIPIEVIKGRHLIFCHSKKCDLAKLVALGNVAY
YGLDVSIVPTSGDVVATDALMTGTGTFDSDVIDCNTCVTVDFSLDPTTETI
TLPDQAVSTORRGRTGKPGIYRFAPGPRSGMFDSSVLCECYDAGCAWELTPA
ETTVRLRAYMTTGLPVCDHLEFWEGVFTGLTHIDAHFLSOTKQSGENPLYLAVQA
TYCARAQPAPPSWDQMKCLIRKPTLHGPTFLLYLGLAVQNEITLTHTPVKYMTCM
SADLEVVVSTWVLVGVLAALAAAYCLSTGCVVIVGRVLSGKPAIIPDREVLRFDE
MEEC"

CDS
63.1%; Score 3713.6; DB 6; Length 6299;

ORIGIN
Query Match

Best Local Similarity 77.4%; Pred. No. 0;		Matches 4876; Conservative 0; Mismatches 1004; Indels 420; Gaps 14;	
QY	2	CGCGCGTTTCGGTGATGACGGTGAACAACTCTGACACATGACGCTCCCGGAGACGGTAC	61
Db	1	CGCGCGTTTCGGTGATGACGGTGAACAACTCTGACACATGACGCTCCCGGAGACGGTAC	60
QY	62	AGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT	121
Db	61	AGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGT	120
QY	122	TGCGGGGTGTGCGGGTGTGCTTAACTATGCGGCATCAGAGCAGATGTACTGAGAGTGCA	181
Db	121	TGCGGGGTGTGCGGGTGTGCTTAACTATGCGGCATCAGAGCAGATGTACTGAGAGTGCA	180
QY	182	CCATATGAAGCTTTTTCGAAAGCCCTAGGCTCCAAAAGAGCCCTCTCACTACTCTCTGGA	241
Db	181	CCATATGAAGCTTTTTCGAAAGCCCTAGGCTCCAAAAGAGCCCTCTCACTACTCTCTGGA	240
QY	242	ATAGCTCAGAGCGCGGCGCTCGGCTCTGTCATAAAATAAAAAATTTAGTCAGCCAT	301
Db	241	ATAGCTCAGAGCGCGGCGCTCGGCTCTGTCATAAAATAAAAAATTTAGTCAGCCAT	300
QY	302	GGGGCGGAGATGGGGCGGAACCTGGGGGGAGGGAATTTATGGCTATTTGGCCATTTGCATA	361
Db	301	GGGGCGGAGATGGGGCGGAACCTGGGGGGAGGGAATTTATGGCTATTTGGCCATTTGCATA	360
QY	362	CGTTGTATCTATATCATATATGATATGATATTAATGCTCATGCTCAATATGACCGCCAT	421
Db	361	CGTTGTATCTATATCATATATGATATTAATGCTCATGCTCAATATGACCGCCAT	420
QY	422	GTTCACATTTGATTTAGCTAGTTATTAATAGTAATCAATTAAGGGGTCAATTTGTCATA	481
Db	421	GTTCACATTTGATTTAGCTAGTTATTAATAGTAATCAATTAAGGGGTCAATTTGTCATA	480
QY	482	GGCCATATATGAGTTCCGGTTTACATAAATTAAGGTAAATGGCCCGCTGGCTGACCGC	541
Db	481	GGCCATATATGAGTTCCGGTTTACATAAATTAAGGTAAATGGCCCGCTGGCTGACCGC	540
QY	542	CCAAACGACCCCGCCCATTTGACGTCAATTAATGACGTATGTTCCCATAGTAAGCCCAATAG	601
Db	541	CCAAACGACCCCGCCCATTTGACGTCAATTAATGACGTATGTTCCCATAGTAAGCCCAATAG	600
QY	602	GGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAATCGCCCACTTGGCAGTAC	661
Db	601	GGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAATCGCCCACTTGGCAGTAC	660
QY	662	ATCAAGTGTATCATATGCCAAGTCCCGCCCTTATTTGACGTCAATGACGTAATGGCCCG	721
Db	661	ATCAAGTGTATCATATGCCAAGTCCCGCCCTTATTTGACGTCAATGACGTAATGGCCCG	720
QY	722	CTTGGCATTTATGCCAGTACATGACCTTACGGGACTTCTCTACTTGGCAGTACATCTACG	781
Db	721	CTTGGCATTTATGCCAGTACATGACCTTACGGGACTTCTCTACTTGGCAGTACATCTACG	780
QY	782	TATTAGTCACTCGTATTATACCATGGTGAATCGGTTTGGCAGTACACCAATGGCGGTGGAT	841
Db	781	TATTAGTCACTCGTATTATACCATGGTGAATCGGTTTGGCAGTACACCAATGGCGGTGGAT	840
QY	842	AGCGGTTTGACTCAACGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGAGTTTGT	901
Db	841	AGCGGTTTGACTCAACGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGAGTTTGT	900
QY	902	TTTGGCACCATAATCAACGGGACTTTCCAAATGTCTAATACCCCGCCCGTTGACGC	961
Db	901	TTTGGCACCATAATCAACGGGACTTTTCCAAATGTCTAATACCCCGCCCGTTGACGC	960
QY	962	AAATGGCGGTAGCGGTGTACGGTGGAGGTCTATATAGCAGAGCTCGTTTAGTGAACC	1021
Db	961	AAATGGCGGTAGCGGTGTACGGTGGAGGTCTATATAGCAGAGCTCGTTTAGTGAACC	1020
QY	1022	GTCCAGTCCGCTGGAGACGCCATTCACGCTGTTTTCACCTCCATAGAACACCCGGGACC	1081
Db			
1021	GTCCAGTCCGCTGGAGACGCCATTCACGCTGTTTTCACCTCCATAGAACACCCGGGACC	1080	
QY	1082	GATCCAGCTCTCGCGCGCGGAAACGGTGCAATTGGAACGGGATTCGCCGTCGCAAGAGTG	1141
Db	1081	GATCCAGCTCTCGCGCGCGGAAACGGTGCAATTGGAACGGGATTCGCCGTCGCAAGAGTG	1140
QY	1142	AGCTAAGTACCGCTATAGACTCTATAGGACACACCCCTTTGGCTCTTATGTCATGCTATAC	1201
Db	1141	AGCTAAGTACCGCTATAGACTCTATAGGACACACACCCCTTTGGCTCTTATGTCATGCTATAC	1200
QY	1202	TGTTTTTGGCTTGGGGCTATACACCCCGCTCTTATGCTATAGTATAGTATAGCTT	1261
Db	1201	TGTTTTTGGCTTGGGGCTATACACCCCGCTCTTATGCTATAGTATAGTATAGCTT	1260
QY	1262	AGCTATAGTGTGGGTATTGACCACTTATGACCACTCCCTTATTTGGTGACCATACTTT	1321
Db	1261	AGCTATAGTGTGGGTATTGACCACTTATGACCACTCCCTTATTTGGTGACCATACTTT	1320
QY	1322	CCATTAATCCATTAACATGGCTCTTTGGCCACAACTATCTCTATTTGGGTATATGCCAAT	1381
Db	1321	CCATTAATCCATTAACATGGCTCTTTGGCCACAACTATCTCTATTTGGGTATATGCCAAT	1380
QY	1382	ACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTACAGGATGGGGTCCATTTAT	1441
Db	1381	ACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTACAGGATGGGGTCCATTTAT	1440
QY	1442	ATTTACAAATTCACATATACAAACCGCTCCCGCTGCCCGCAGTCTTTTATTAACAT	1501
Db	1441	ATTTACAAATTCACATATACAAACCGCTCCCGCTGCCCGCAGTCTTTTATTAACAT	1500
QY	1502	AGCGTGGGATCTCCGACATCTCGGTPACGTGTTCCGGAATGGGTCTTCTCCGGTAGCG	1561
Db	1501	AGCGTGGGATCTCCGACATCTCGGTPACGTGTTCCGGAATGGGTCTTCTCCGGTAGCG	1560
QY	1562	GGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCGAGCGCTCATGTCGCTCGGCA	1621
Db	1561	GGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCGAGCGCTCATGTCGCTCGGCA	1620
QY	1622	GCTCTTGTCTCTAACAGTGGAGGCGCAGACTTAGGCACAGCAATGCCACCAACACCA	1681
Db	1621	GCTCTTGTCTCTAACAGTGGAGGCGCAGACTTAGGCACAGCAATGCCACCAACACCA	1680
QY	1682	GTGTCGCGCAAGGCGGTGGGGTAGGTATGTGTCTGAAATAGCTCGGAGATTTGG	1741
Db	1681	GTGTCGCGCAAGGCGGTGGGGTAGGTATGTGTCTGAAATAGCTCGGAGATTTGG	1740
QY	1742	CTCGCACCTGGAGCGAGATGGAAGACTTAAGCGCGGAGAGAGATGCGAGCAGCTG	1801
Db	1741	CTCGCACCTGGAGCGAGATGGAAGACTTAAGCGCGGAGAGAGATGCGAGCAGCTG	1800
QY	1802	AGTTGTTGTATTCTGATAAGAGTCAGAGTAACCTCCCGTTGCGGTCTGTTAAACGGTGA	1861
Db	1801	AGTTGTTGTATTCTGATAAGAGTCAGAGTAACCTCCCGTTGCGGTCTGTTAAACGGTGA	1860
QY	1862	GGGCACTGTAGTCTGAGCAGTACTGTTGCTGCCGCGCGGCCACAGACATAATAGCTG	1921
Db	1861	GGGCACTGTAGTCTGAGCAGTACTGTTGCTGCCGCGCGGCCACAGACATAATAGCTG	1920
QY	1922	ACAGACTAACAGACTGTTCTTTCCATGGTCTTTTCTGACGTACCGCTCGTGCAGCAAT	1981
Db	1921	ACAGACTAACAGACTGTTCTTTCCATGGTCTTTTCTGACGTACCGCTCGTGCAGCAAT	1980
QY	1982	TCAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTCTG-----	2027
Db	1981	GAATTCACCATGGCGCCCATCAGCGGTACGCCAGCACAAGGGGCTCTTAGGGTGC	2040
QY	2028	-----CTGCTGTGTGGAGCAGTCTTCGTTTCG	2054
Db	2041	ATAATCACCAGCTTAATGGCGGGACAAAAACCAAGTGGAGGTGAGGTCCAGATTGTG	2100
QY	2055	CCAGCGCTTAGCAAC-----C	2072
Db	2101	TCAACTGCTGCCAAAACCTTCTCGCAACGTGCAATCAATGGGGTGTGCTGGAGTGTCTAC	2160

|||||
Db 4320 ACTCGTGGCGTCCGTCGTCGTCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA 4379
|||||
QY 3963 TACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTCAGCAAAAGGCCAGC 4022
|||||
Db 4380 TACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAGAACATGTCAGCAAAAGGCCAGC 4439
|||||
QY 4023 AAAAGCCAGGAACCGCTAAAGAGCCGCGTGTGTCGGCGTTTTTCCATAGGCTCCGCCCC 4082
|||||
Db 4440 AAAAGCCAGGAACCGCTAAAGAGCCGCGTGTGTCGGCGTTTTTCCATAGGCTCCGCCCC 4499
|||||
QY 4083 CTGACAGCATACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT 4142
|||||
Db 4500 CTGACAGCATACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT 4559
|||||
QY 4143 AAAGATACAGGGGTTTTCCCGCTGGAGGCTCCCTCGTCGGCTCTCTGTTCCGACCCCTGC 4202
|||||
Db 4560 AAAGATACAGGGGTTTTCCCGCTGGAGGCTCCCTCGTCGGCTCTCTGTTCCGACCCCTGC 4619
|||||
QY 4203 CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCAATGCT 4262
|||||
Db 4620 CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAAGCGTGGCGCTTTCTCAATGCT 4679
|||||
QY 4263 CAGCTGTAGGTATCTCAGTTCGGTGTAGTCTGTCCTTCCGCTTCAAGCTGGGCTGTGTGCA 4322
|||||
Db 4680 CAGCTGTAGGTATCTCAGTTCGGTGTAGTCTGTCCTTCCGCTTCAAGCTGGGCTGTGTGCA 4739
|||||
QY 4323 AACCCCGCTTCAGCCGCGCGCTGCGCTTATCCGCTAACCTATCGCTTGGTTCAGTCCAA 4382
|||||
Db 4740 AACCCCGCTTCAGCCGCGCGCTGCGCTTATCCGCTAACCTATCGCTTGGTTCAGTCCAA 4799
|||||
QY 4383 CGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCTGCTGAACAGGATTAGCAGAGCA 4442
|||||
Db 4800 CGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCTGCTGAACAGGATTAGCAGAGCA 4859
|||||
QY 4443 GGTATCTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAA 4502
|||||
Db 4860 GGTATCTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAA 4919
|||||
QY 4503 GGACAGTATTTGGTATCTCGCTCTCTGCTGAAGCAGTACCTTCGGAAGAGTTGGTA 4562
|||||
Db 4920 GGACAGTATTTGGTATCTCGCTCTCTGCTGAAGCAGTACCTTCGGAAGAGTTGGTA 4979
|||||
QY 4563 GCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGC 4622
|||||
Db 4980 GCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGC 5039
|||||
QY 4623 AGATTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTG 4682
|||||
Db 5040 AGATTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTG 5099
|||||
QY 4683 AGCTCAGTGGAGAAACCTCAGTGAAGGATTTTGGTCATGAGATTATCAAAAGGA 4742
|||||
Db 5100 AGCTCAGTGGAGAAACCTCAGTGAAGGATTTTGGTCATGAGATTATCAAAAGGA 5159
|||||
QY 4743 TCTTACCTAGATCTTTTAAATTAAGATTTAAATTAAGATTTAAATTAAGATTTAAATTAAG 4802
|||||
Db 5160 TCTTACCTAGATCTTTTAAATTAAGATTTAAATTAAGATTTAAATTAAGATTTAAATTAAG 5219
|||||
QY 4803 AGTAAACTTGGTCTGACGTTACCAATGCTTAATCAGTAGGAGCCTATCTCAGCGATCT 4862
|||||
Db 5220 AGTAAACTTGGTCTGACGTTACCAATGCTTAATCAGTAGGAGCCTATCTCAGCGATCT 5279
|||||
QY 4863 GTCTATTTCGTTTCATCATAGTTCGCTGACTCCCGTCTGTGTAGATACTACGATACGG 4922
|||||
Db 5280 GTCTATTTCGTTTCATCATAGTTCGCTGACTCCCGTCTGTGTAGATACTACGATACGG 5339
|||||
QY 4923 AGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCGGCTC 4982
|||||
Db 5340 AGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCGGCTC 5399
|||||
QY 4983 CAGATTATCAGCAATAAACCGACCGCGGAAGGGCCGAGCGCAGAGTGTCTTGCAA 5042
|||||

Db 5400 CAGATTATCAGCAATAAAACCGACCGCCGAGGCGGAGGCGCAGAGTGTGCTTGCAA 5459
QY 5043 CTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTGTGCG 5102
Db 5460 CTTTATCCGCCCTCCATCCAGTCTATTAAATTTGTCGGGAGAGCTAGAGTAAGTGTGCG 5519
QY 5103 CAGTTAATAAGTTTGGCAAACGTTGTCATTTGCTACAGGCATCGTGGTGTCAAGTGTG 5162
Db 5520 CAGTTAATAAGTTTGGCAAACGTTGTCATTTGCTACAGGCATCGTGGTGTCAAGTGTG 5579
QY 5163 CGTTTGGTATGCTTCAATTCAGTCCGGTCCCAAGCATCAAGCGAGTTCATGATGCC 5222
Db 5580 CGTTTGGTATGCTTCAATTCAGTCCGGTCCCAAGCATCAAGCGAGTTCATGATGCC 5639
QY 5223 CCATGTTGTGCAAAAAGCGGTTAGCTCTCGTCTCCGATCGTGTGTCAGAAAGTAA 5282
Db 5640 CCATGTTGTGCAAAAAGCGGTTAGCTCTCGTCTCCGATCGTGTGTCAGAAAGTAA 5699
QY 5283 TGGCCGAGTGTATCACTCATGTTTATGCGACACTGCATATAATCTTCTTACTGTATGC 5342
Db 5700 TGGCCGAGTGTATCACTCATGTTTATGCGACACTGCATATAATCTTCTTACTGTATGC 5759
QY 5343 CATCCGTAAGATGCTTTCTGTCAGTGTGAGTACTCAACCAAGTCAATCTGTGAGTATG 5402
Db 5760 CATCCGTAAGATGCTTTCTGTCAGTGTGAGTACTCAACCAAGTCAATCTGTGAGTATG 5819
QY 5403 GTATGCGGCGACCGAGTGTCTCTTGGCCGCGCTCAATACGGGATAATACCGCCGACATA 5462
Db 5820 GTATGCGGCGACCGAGTGTCTCTTGGCCGCGCTCAATACGGGATAATACCGCCGACATA 5879
QY 5463 GCAGAACTTTAAAGTGTCTCATTTGGAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGA 5522
Db 5880 GCAGAACTTTAAAGTGTCTCATTTGGAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGA 5939
QY 5523 TCTTACCGCTGTGTGAGATCCAGTTCGATTAACCACTCGTCGACCACTGATCTTCAG 5582
Db 5940 TCTTACCGCTGTGTGAGATCCAGTTCGATTAACCACTCGTCGACCACTGATCTTCAG 5999
QY 5583 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAATGCCGCAA 5642
Db 6000 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAATGCCGCAA 6059
QY 5643 AAAAGGGAATAAGGGCGACACGGAATTTGAATACTCATACTCTTCTTCTTTTCAATATT 5702
Db 6060 AAAAGGGAATAAGGGCGACACGGAATTTGAATACTCATACTCTTCTTCTTTTCAATATT 6119
QY 5703 ATTGAAGCAATTATCAGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGA 5762
Db 6120 ATTGAAGCAATTATCAGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGA 6179
QY 5763 AAAATAAACAATAAGGGTTTCCGCGCACATTTCCCGAAAAAGTCCACCTGAGCTCTAAG 5822
Db 6180 AAAATAAACAATAAGGGTTTCCGCGCACATTTCCCGAAAAAGTCCACCTGAGCTCTAAG 6239
QY 5823 AAACCATTTATCATGACATTAACCTATAAAATAAGGGGTATCAGAGGCCCTTTTCGTC 5882
Db 6240 AAACCATTTATCATGACATTAACCTATAAAATAAGGGGTATCAGAGGCCCTTTTCGTC 6299

RESULT 5
AF286076
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AF286076 5130 bp DNA circular SYN 27-DEC-2000
Expression vector AF286076, complete sequence.

AF286076 GI:12034709

Expression vector AF286076
Expression vector AF286076
other sequences: artificial
sequences; vectors.
1 (bases 1 to 5130)
Arthos,J. and Mullins,J.I.
Multipurpose Eukaryotic Expression Vector
Unpublished

1566 AGCTTCCACATCCGAGCCCTGGTCCATCCGTCAGCGGCTCATGTCGCTCGGCGAGTC 1625
1611 AGCTTCCACATCCGAGCCCTGGTCCATCCGTCAGCGGCTCATGTCGCTCGGCGAGTC 1670
1626 CTTGCTCCTTAACAGTGGAGCCAGACTTAGGCAAGCAATCCACCAACCAAGTGT 1685
1671 CTTGCTCCTTAACAGTGGAGCCAGACTTAGGCAAGCAATCCACCAACCAAGTGT 1730
1686 GCCGCAACAGGCCGCTGGCGTATGCTGTGAAATAGCTCGGAGATTGGGCTCG 1745
1731 GCCGCAACAGGCCGCTGGCGTATGCTGTGAAATAGCTCGGAGATTGGGCTCG 1790
1746 CACTCGACGACATGGAAGCTTAAGGCAAGCGGCAAGAAAGATCGAGCAGCTCAGTT 1805
1791 CACCGTGAACGACATGAAGACTTAAGGCAAGCGGCAAGAAAGATCGAGCAGCTCAGTT 1850
1806 GTTGTATTCTGATAAGAGTCAGAGTTAACTCCGTTGCGGTGCTGTAAAGGAGGCG 1865
1851 GTTGTATTCTGATAAGAGTCAGAGTTAACTCCGTTGCGGTGCTGTAAAGGAGGCG 1910
1866 AGTGTAGTCTGAGCAGTACTGCTGTGCGCGCGCCACCAAGACATATAGCTGACAG 1925
1911 AGTGTAGTCTGAGCAGTACTGCTGTGCGCGCGCCACCAAGACATATAGCTGACAG 1970
1926 ACTAAGCAGACTGTTCTCTTTTCCATGAGTCTTTCTGCAAGTCAACGCTCGTCAAGTCA 1985
1971 ACTAAGCAGACTGTTCTCTTTTCCATGAGTCTTTCTGCAAGTCAACGCTCGTCAAGTCA 2023
1986 GCATCATGATGCAATGAAGAGAGGCTCTGCTGTGCTGTGCTGTGCTGTGAGGAGCTC 2045
2024 GCAATCATGATGCAATGAAGAGAGGCTCTGCTGTGCTGTGCTGTGCTGTGAGGAGCTC 2083
2046 TTGCTTTTCCGCAAGCTAGCGAAACCCAGTCAACCGGGGAGTCCCGCCACACTGTG 2105
2084 TTGCTTTTCCGCTAGCGCGG-----ACGGTGGATCTCTGCAATCCCTA 2127
2106 TCTGGATTTTGTAGCTCTCTCGCACGAGCGCAAGCAGACGTCAGCTGATCAACACC 2165
2128 GGAGGATTAG-----GC 2139
2166 AACGGCAGTTGGACCTCAATAGCAGCGGCTGAACCTGCAATGATAGCTCAACACCGGC 2225
2140 AAGGGCTTGGAGCTCAGCTCTTGTGAGGAGCAGAAATACAAATCAGGGGAGTATATGAAT 2199
2226 TGGTTTGGCAGGCTTTTCTATACCAACAAAGTTCAACTCTTCAAGGCTCTCTGAGAGGCTA 2285
2200 ACTCCATGAGAAACCCAGATCTACGATATGATCAGCTCTGCTGCTGCTCTCTAGTTGCCA 2259
2286 GCCAGCTGCGGACCCCTTACCGAATTTTGACAGGGCTGGGGCCCTATCAGTTATGCCAAC 2345
2260 GCCATCTGTTGTTGCCCCCTC-----CCCCGTGCTTCTTGAACCT 2301
2346 GGNAGCGGCCCGACAGCGCCCTACTGCTGACACTACCCCCCAAAACCTTGGCGTAT 2405
2302 GGAAGTGCCACTCCCACTGCTCT-----TCCTAATAAAATGAGGAAAT 2346
2406 GTGCCGCGAAGAGTGTGTGCTGCTGCTATATTGCTTCACTCCAGCCCCGCTGGTGG 2465
2347 TGCATCGCATTTCTGAGTAGGTGTCACTTAT-----TCTGGGGGGTGGGGTGGG 2398
2466 GGAACGACCGACAGGTGCGGCGCGCCACCTACAGCTGGGGTGAATAATGATAGCGACGTC 2525
2399 CAGGACAGCAAGGGGAGGATTGGGAAGACATAGCAGGCATGCTGGGGATGCGG----- 2453
2526 TTGCTCCTTAACAAATACAGGCCACCGCTGGGCAATTTGGTTGCGTTGTACCTGGAATGAAC 2585
2454 -----TGGGCTCTATGGAACCAAGCTGGGG 2478
2586 TCAACTGGATTACCAAAGTGTGGGAGCGGCTCTTGTGTCTCATCGGAGGGGCGGCAC 2645
2479 TCGACAGGA-----CCGCTATC 2495
2646 AACACCCCTGCACTGCCCACTGATTGCTTCCGCAAGCATCCGGAACCCACATACCTCTCGG 2705

2496 AGGACATAGCTGGTACCGTGATATTGCTGAAGAGCTTGGCGGCA----- 2544
2706 TGGGGCTCGGTCCTCGGATCACACCCAGTGCCTGGTGCAGCTACCCGCTATAGCTTTGG 2765
2545 -----ATGGGCTGACCGCTTCTCGTCTTTACGGTATCGGCGCTCCCGATTGCG 2594
2766 CATTATCTTTGACCACTCAACTACACATATTTAAAATCAGGATGTACGTTGGAGGGGTC 2825
2595 AGGCAATCGCTTCTATCGCTTCTTGACGAGTTCTTCTGAGCGGAGCTCTGGGTTTC-- 2652
2826 GAACACAGGCTGGAAGCTCGCTGCAACTGGACCGGGGGCAACGTTGGCGATCTGGAAGAT 2885
2653 ----- 2652
2886 AGGACAGCTCCGAGATCGATATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTC 2945
2653 ---GAAATCAGCGACCAAGCGAGCGCCAACTGCCATCAGAGATTTCGATTCCACGCGC 2709
2946 GTGTTACAGCGCGGGTTTTCTTTGTTGACAAGAACTCTCAATACCGCAGAGTCTAGAC 3005
2710 GCCTTCTATGAAGGTTGGGCTTCGAA----- 2737
3006 TCGTGGTGAAGTCTCTCAATTTTCTAGGGGATCTCCCGTGTGCTTGGCCAAAATTCG 3065
2738 ----- 2737
3066 CAGTCCCCAACCTCCNAATCACTCACCAACCTCTCTGCTCCCAATTTGTCTGCTTATGCG 3125
2738 -----TCGTTTTTCGGGAGCGCGGC 2757
3126 TGGATGTGTCTGGGGGTTTTATCTCTTCACTCTGCTGTGCTATGCTCTCATCTTC 3185
2758 TGGATGATCTCCAGCGCGGGGATC-----TCATGCTGGAGTCTTTCGCCACCCC 2808
3186 TTATTGGTTCTTCTGAGATTATCAAGGTATGTTGCCGTTTGTCTCTAATTTCCAGGATCA 3245
2809 AACTGTTTATGTCAGCTTATAATGG----- 2834
3246 ACNACCAACGATACGGGACCATGCAAAACCTGACAGACTCTCTGCTCAAGGCAACTCTATG 3305
2835 ----- 2834
3306 TTTTCCCTCATGTTGCTGTACAAAACCTCAGGATGGAATTTGACCTGTATTTCCCATCCCA 3365
2835 ---TTACAAATAAGCAATAGCATCAAAATTTACAAATAAAGCAATTTTTCCTGCTAT 2892
3366 TCGTCTGGGCTTTTCGCAAAATACCTATGAGGAGTGGGCTCAGTCCGTTTCTCTGGCTC 3425
2893 TCTAGTTGTGGTTTGTGCCAA----- 2913
3426 AGTTTACTAGTGCCATTGTTTCAAGTGGTTCGTTAGGGCTTTTCCCCCACTGTTGGCTTTCA 3485
2914 -----CTCA 2917
3486 GCTATATGATGATGTGGTATTTGGGGCCAAAGTCTGTACAGCATCGTAGTCCCTTTATA 3545
2918 TCAATGATCTTATCATGTCGATCGCGCGCGGATCCGTCGAGAGCTTGGCGTAAATC 2977
3546 CGGCTGTTACCAATTTTCTTCTCTGGGTATACATTTAAAGAAATTCAGACTCGAGCAA 3605
2978 ATGGTCATAGCTGTTTCTCTGTGTGAAATTTGTTATCGCTCACAATTTCCACACAATACG 3037
3606 GTCTAGAAAGGCGCGCAAGATATCAAGGATCCACTACCGGTTAGAGCTCGGTGATCAGC 3665
3038 AGCCGGAAGCAATAAGTGTAAAGCCCTGGGGTGCCTAATGAGTGAGTAACTCAATTAAT 3097
3666 CTCGACTGTGCTTCTAGTTGCGAGGCACTGTGTTGCTCCCTCCCTCCCGTCTCTCTT 3725
3098 TCGG-----TTGGCTCACTGCCCGCTTTC-- 3123
3726 GACCCTGGAAAGTGCCACTCCCACTGTCTTCTTAATAAAATGAGGAAATTCGATCGCA 3785

Db 3124 ----- 3123
QY 3786 TTGTCAGTAGGTGTCATCTTATTTCTGGGGGGTGGGGTGGGCGCAGGACAGCAAGGGGA 3845
Db 3124 -AGTCGGGAAACCTGTCGTGCCAGCTGCATTTATGAATCGGCCAACCGCGGGGAGAGGC 3182
QY 3846 GGATTCCGGAAGACAAATAGCAGGCATGCTGGGAGCTCTCCGGCTCCTCGCTCACTGACT 3905
Db 3183 GGTTCGCGTA-----TTGGGCGCTCTCCGGCTCCTCGCTCACTGACT 3225
QY 3906 CGTCGGCTCGGTTCGGTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAATAC 3965
Db 3226 CGTCGCTCGGTTCGGTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAATAC 3285
QY 3966 GGTATTCACAGAAATCAGGGGATAACGACAGGAAGAAACATGTGAGCAAAAGGCGCAGCAA 4025
Db 3286 GGTATTCACAGAAATCAGGGGATAACGACAGGAAGAAACATGTGAGCAAAAGGCGCAGCAA 3345
QY 4026 AGGCCAGGAACCGTAAAGAGCGCGTGTCTGCGCTTTTCCATAGGCTCCGCCCCCTG 4085
Db 3346 AGGCCAGGAACCGTAAAGAGCGCGTGTCTGCGCTTTTCCATAGGCTCCGCCCCCTG 3405
QY 4086 ACAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 4145
Db 3406 ACAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAA 3465
QY 4146 GATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCCTCGCG 4205
Db 3466 GATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCCTCGCG 3525
QY 4206 TTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAC 4265
Db 3526 TTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAC 3585
QY 4266 GCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAGCTGGGCTGTGTGCAAGAAC 4325
Db 3586 GCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAGCTGGGCTGTGTGCAAGAAC 3645
QY 4326 CCCCCGTTGAGCCGACCGCTGGCGCTTATCCGGTAACTATCGCTTCGAGTCCAAACCCGG 4385
Db 3646 CCCCCGTTGAGCCGACCGCTGGCGCTTATCCGGTAACTATCGCTTCGAGTCCAAACCCGG 3705
QY 4386 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCCAGGT 4445
Db 3706 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCCAGGT 3765
QY 4446 ATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGCTAACTACGGCTACACTAGAGGA 4505
Db 3766 ATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGCTAACTACGGCTACACTAGAGGA 3825
QY 4506 CAGTATTTGGTATCTGGCTCTGCTGAAGCAGTTACCTTCGGAAAAAGATTGGTAGCT 4565
Db 3826 CAGTATTTGGTATCTGGCTCTGCTGAAGCAGTTACCTTCGGAAAAAGATTGGTAGCT 3885
QY 4566 CTTGATTCGGCAACAAACACCGCTGTAGCGGTGTTTTTTTGTTCGAACGACGAGA 4625
Db 3886 CTTGATTCGGCAACAAACACCGCTGTAGCGGTGTTTTTTTGTTCGAACGACGAGA 3945
QY 4626 TTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTTCTACGGGCTGACG 4685
Db 3946 TTACGGCGAGAAAAAAGGATCTCAAGAGATCCTTTTGATCTTTTCTACGGGCTGACG 4005
QY 4686 CTCAGTGGAAACGAAAACTACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 4745
Db 4006 CTCAGTGGAAACGAAAACTACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 4065
QY 4746 TCACCTAGATCCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATATATAGT 4805
Db 4066 TCACCTAGATCCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATATATAGT 4125
QY 4806 AAACCTTGGTGTACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTC 4865
Db 4126 AAACCTTGGTGTACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTC 4185

QY 4866 TATTTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAACTACGATACGGAGG 4925
Db 4186 TATTTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAACTACGATACGGAGG 4245
QY 4926 GCTTACCATCTCGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAG 4985
Db 4246 GCTTACCATCTCGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAG 4305
QY 4986 ATTTATCAGCAATAAACACAGCCAGCGGAAAGGGCCGAGCGCAGAGAGTGTCTGCAACTT 5045
Db 4306 ATTTATCAGCAATAAACACAGCCAGCGGAAAGGGCCGAGCGCAGAGTGTCTGCAACTT 4365
QY 5046 TATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAGTAAGTGTTCGCCAG 5105
Db 4366 TATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAGTAAGTGTTCGCCAG 4425
QY 5106 TTAATAGTTTGGCGCAACGTTGTTGTCATGCTACAGGCATCGTGTGTCAACGCTCGT 5165
Db 4426 TTAATAGTTTGGCGCAACGTTGTTGTCATGCTACAGGCATCGTGTGTCAACGCTCGT 4485
QY 5166 TTGGTATGCTTCATTCAGCTCCGTTCCAAACGATCAAGGCGAGTTACATGATCCCA 5225
Db 4486 TTGGTATGCTTCATTCAGCTCCGTTCCAAACGATCAAGGCGAGTTACATGATCCCA 4545
QY 5226 TGTGTGCAAAAAAGCGGTAGTCTCTTCGCTCCGATCGTGTGTCAGAGTAAGTTGG 5285
Db 4546 TGTGTGCAAAAAAGCGGTAGTCTCTTCGCTCCGATCGTGTGTCAGAGTAAGTTGG 4605
QY 5286 CCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTCCTTTACTGTCAATGCCAT 5345
Db 4606 CCGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTCCTTTACTGTCAATGCCAT 4665
QY 5346 CCGTAAGATGCTTTTCTGTGACTGTGTGACTCTCAACCAAGTCAATCTGAGAATAGTGA 5405
Db 4666 CCGTAAGATGCTTTTCTGTGACTGTGTGACTCTCAACCAAGTCAATCTGAGAATAGTGA 4725
QY 5406 TCGGGCGACCGAGTTGCTCTTCGCGCGCTCAATACGGGATAATACCGGCGCACATAGCA 5465
Db 4726 TCGGGCGACCGAGTTGCTCTTCGCGCGCTCAATACGGGATAATACCGGCGCACATAGCA 4785
QY 5466 GAACTTTAAAGTGTCTCATCTTTGAAAAACGTTCTTCGGGGGAAAACTCTCAAGGATCT 5525
Db 4786 GAACTTTAAAGTGTCTCATCTTTGAAAAACGTTCTTCGGGGGAAAACTCTCAAGGATCT 4845
QY 5526 TACCGCTGTGTGAGATCCAGTTCCGATGTAACCACTCGTGCAACCACTGATCTTCAGCAT 5585
Db 4846 TACCGCTGTGTGAGATCCAGTTCCGATGTAACCACTCGTGCAACCACTGATCTTCAGCAT 4905
QY 5586 CTTTACTTTTCAACGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAA 5645
Db 4906 CTTTACTTTTCAACGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAA 4965
QY 5646 AGGGAATAAGGCGCAGCACGAAATGTTGAATACTCATACTCTTTCTTTTCAATATATT 5705
Db 4966 AGGGAATAAGGCGCAGCACGAAATGTTGAATACTCATACTCTTTCTTTTCAATATATT 5025
QY 5706 GAAGCAATTTACAGGGTTATTTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAA 5765
Db 5026 GAAGCAATTTACAGGGTTATTTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAA 5085
QY 5766 ATAAACAAATAGGGTTCCGCGCACATTTCCCGCAAAAGTGCAC 5810
Db 5086 ATAAACAAATAGGGTTCCGCGCACATTTCCCGCAAAAGTGCAC 5130

RESULT 6
AX154496
LOCUS
DEFINITION
AX154496
ACCESSION
AX154496.1
GI:14536106
KEYWORDS

AX154496 Sequence 1 from Patent W00138358.
AX154496
AX154496.1 GI:14536106

linear

4276 bp. DNA

PAT 22-JUN-2001

SOURCE	synthetic construct			
ORGANISM	synthetic construct			
	other sequences; artificial sequences.			
REFERENCE	1			
AUTHORS	Selby, M.C., Glazer, B.C. and Houghton, M.C.			
TITLE	Hdv/hcv virus-like particle			
JOURNAL	Patent: WO 0138358-A 1 31-MAY-2001;			
	CHIRON CORPORATION (US)			
FEATURES	Location/Qualifiers			
source	1..4276			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="plasmid pCMVII"			
ORIGIN	Query Match 45.2%; Score 2658.4; DB 6; Length 4276;			
	Best Local Similarity 72.7%; Pred. No. 0;			
	Matches 4275; Conservative 0; Mismatches 1; Indels 1606; Gaps 1;			
QY	1	TCGCGGCTTCGGGTGATGACGGTGAAACCTCTGACACATGACAGCTCCCGGAGACGGTCA	60	
DB	1	TCGCGGCTTCGGGTGATGACGGTGAAACCTCTGACACATGACAGCTCCCGGAGACGGTCA	60	
QY	61	CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGGTCAAGCGGGTG	120	
DB	61	CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGGTCAAGCGGGTG	120	
QY	121	TTGGCGGGTTCGGGCTGCTTAACTATGCGGCATCAGAGCAGATGTTACTGAGAGTGC	180	
DB	121	TTGGCGGGTTCGGGCTGCTTAACTATGCGGCATCAGAGCAGATGTTACTGAGAGTGC	180	
QY	181	ACCATATGAAGCTTTTGTGAAAAGCCTAGGCCCTCAAAAAGCCTCTCACTACTTCTGG	240	
DB	181	ACCATATGAAGCTTTTGTGAAAAGCCTAGGCCCTCAAAAAGCCTCTCACTACTTCTGG	240	
QY	241	AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAATAAAAAAATTAAGTCAGCCA	300	
DB	241	AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAATAAAAAAATTAAGTCAGCCA	300	
QY	301	TGGGCGGAGAAATGGCGGAACTGGGCGGGAGGGAATATTGGCTATTGGCCATTGCAT	360	
DB	301	TGGGCGGAGAAATGGCGGAACTGGGCGGGAGGGAATATTGGCTATTGGCCATTGCAT	360	
QY	361	ACGTTGTATCATATCATATATGATATTAATTGGCTCATGTCCAATATGACCGCCA	420	
DB	361	ACGTTGTATCATATCATATATGATATTAATTGGCTCATGTCCAATATGACCGCCA	420	
QY	421	TGTTGACATTGATTTAGCTAGTTAATATAGTAATCAATTACGGGGTCATTAGTTCAAT	480	
DB	421	TGTTGACATTGATTTAGCTAGTTAATATAGTAATCAATTACGGGGTCATTAGTTCAAT	480	
QY	481	AGCCCATATAGAGTTCCGGTTACATACTTACGGTAATGGCGCGCTGGCTGACCG	540	
DB	481	AGCCCATATAGAGTTCCGGTTACATACTTACGGTAATGGCGCGCTGGCTGACCG	540	
QY	541	CCCAACGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAAGCCCAATA	600	
DB	541	CCCAACGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAAGCCCAATA	600	
QY	601	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCATTTGGCAGTA	660	
DB	601	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCATTTGGCAGTA	660	
QY	661	CATCAAGTGTATCATATGCAAGTCCGCCCTTATTGACGTCAATGACGGTAATGGCCC	720	
DB	661	CATCAAGTGTATCATATGCAAGTCCGCCCTTATTGACGTCAATGACGGTAATGGCCC	720	
QY	721	GCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC	780	
DB	721	GCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC	780	
QY	781	GTATTAGTCATCGCTATTACCATGGTGTGCGGTTTTTGGCAGTACACCAATGGCGGTGA	840	
781		GTATTAGTCATCGCTATTACCATGGTGTGCGGTTTTTGGCAGTACACCAATGGCGGTGA	840	
841		TAGCGGTTTGACCTCACGGGATTTCCAAAGTCTCCACCCCATTCACCTCAATGGAGTTTG	900	
841		TAGCGGTTTGACCTCACGGGATTTCCAAAGTCTCCACCCCATTCACCTCAATGGAGTTTG	900	
901		TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCTGAATAACCCCGCCGTTGACG	960	
901		TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCTGAATAACCCCGCCGTTGACG	960	
961		CAAATGGGCGGTAGGCGGTGACGGTGGAGTCTATATAAGCAGAGTCTGTTTGTAGTAAC	1020	
961		CAAATGGGCGGTAGGCGGTGACGGTGGAGTCTATATAAGCAGAGTCTGTTTGTAGTAAC	1020	
1021		CGTCAGATCGCTGGAGAGCCCATCCACGCTGTTTGTGACCTCCATAGACACACCGGAC	1080	
1021		CGTCAGATCGCTGGAGAGCCCATCCACGCTGTTTGTGACCTCCATAGACACACCGGAC	1080	
1081		CGATCCAGCTCCGCGCGGGAAACGGTGCATTTGGAAACGGATTCCTCCGTGCAAGAGT	1140	
1081		CGATCCAGCTCCGCGCGGGAAACGGTGCATTTGGAAACGGATTCCTCCGTGCAAGAGT	1140	
1141		GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGATA	1200	
1141		GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGATA	1200	
1201		CTGTTTTCGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT	1260	
1201		CTGTTTTCGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT	1260	
1261		TAGCCTATAGGTGTTGTTTATGACCACTTATGACCACTCCCCCTATTGGTGACGATCTT	1320	
1261		TAGCCTATAGGTGTTGTTTATGACCACTTATGACCACTCCCCCTATTGGTGACGATCTT	1320	
1321		TCATTATCAATCCATAAACATGCTCTTTGGCCAACTATCTCTATGCTATATGCTTAT	1380	
1321		TCATTATCAATCCATAAACATGCTCTTTGGCCAACTATCTCTATGCTATATGCTTAT	1380	
1381		TACTCTGCTCTCAGAGCTGACAGGACTCTGTTATTTTACAGGATGGGTCCATTTAT	1440	
1381		TACTCTGCTCTCAGAGCTGACAGGACTCTGTTATTTTACAGGATGGGTCCATTTAT	1440	
1441		TATTTACAAATTCACATATACAAACGCGCTCCCGCGCCGAGTTTATTAACA	1500	
1441		TATTTACAAATTCACATATACAAACGCGCTCCCGCGCCGAGTTTATTAACA	1500	
1501		TAGCGTGGATCTCCGACATCTCGGTAAGTCTCGGACATGCGGCTCTTCTCCGCTAGC	1560	
1501		TAGCGTGGATCTCCGACATCTCGGTAAGTCTCGGACATGCGGCTCTTCTCCGCTAGC	1560	
1561		GGGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCCAGCGGCTCATGGTCTCGGC	1620	
1561		GGGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCCAGCGGCTCATGGTCTCGGC	1620	
1621		AGCTCTCTCTCTTAACAGTGGAGCCAGACTTTAGGCACAGACAAATGCCACACACC	1680	
1621		AGCTCTCTCTCTTAACAGTGGAGCCAGACTTTAGGCACAGACAAATGCCACACACC	1680	
1681		AGTGTGCCCAAAAGCCCTGGCGGTAGGTATGTCTGTAATAATGAGCTCGAGATGG	1740	
1681		AGTGTGCCCAAAAGCCCTGGCGGTAGGTATGTCTGTAATAATGAGCTCGAGATGG	1740	
1741		GCTCGACCTGACAGAGTGAAGTAAAGGACGGCAGAGAGATGACAGGAGCT	1800	
1741		GCTCGACCTGACAGAGTGAAGTAAAGGACGGCAGAGAGATGACAGGAGCT	1800	
1801		GAGTTGTGTATTTCTGATAAGAGTACAGAGTAACTCCCGCTGGGTGCTGTTAAACGGTG	1860	
1801		GAGTTGTGTATTTCTGATAAGAGTACAGAGTAACTCCCGCTGGGTGCTGTTAAACGGTG	1860	
1861		AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCCGCGGCCACACAGACATAATAGCT	1920	

Db 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTCTCGCGCGCGGCCACGACGACATAATAGCT 1920
Qy 1921 GACAGACTAACAGACTGTCTCTTTCCATGGGTCCTTTTCTGCAGTCAACCGTGTGACGAA 1980
Db 1921 GACAGACTAACAGACTGTCTCTTTCCATGGGTCCTTTTCTGCAGTCAACCGTGTGACGAA 1977
Qy 1981 TTCAAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGTGCTGTGTGGAG 2040
Db 1978 ----- 1977
Qy 2041 CAGTCTTCGTTTCGCCCGAGGCTAGCGAAACCCACGTCACCGGGGGAAGTGCOCGCCACA 2100
Db 1978 ----- 1977
Qy 2101 CTGTGTCTGGATTGTAGCCTCTCCGACAGGCGCCCTGAACCTGCAATGATAGCCTCAACA 2160
Db 1978 ----- 1977
Qy 2161 ACACCAACGGCAGTTGGCACCCTCAATAGCACGGCCCTGAACCTGCAATGATAGCCTCAACA 2220
Db 1978 ----- 1977
Qy 2221 CCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCTGAGA 2280
Db 1978 ----- 1977
Qy 2281 GGCTAGCCAGCTGCCAGCCCTTACCGATTTCGATTTTGACGAGGCTGGGGCCCTATCAGTTATG 2340
Db 1978 ----- 1977
Qy 2341 CCNACGGAAGCGCCCGACCGGCCCTACTGTGGCACTACCCCCCAAAACCTTGGC 2400
Db 1978 ----- 1977
Qy 2401 GTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTGCTTCACTCCAGCGCCGTG 2460
Db 1978 ----- 1977
Qy 2461 TGGTGGGAACGACCGACAGGTCCGGCGCGCCACCTACAGCTGGGTGAAATGATACGG 2520
Db 1978 ----- 1977
Qy 2521 ACGTCTTCGTCTTAACATACAGCGCCACCGCTGGGCAATTGGTTGGTTGTACCTGGA 2580
Db 1978 ----- 1977
Qy 2581 TGAACTCAACTGGAATTCACCAAGTGTGCGAGCGCTCCTTGTGTATCGAGGGCGG 2640
Db 1978 ----- 1977
Qy 2641 GCAACAACACCTGCACCTGCCCACTGATTTCCGCAAGCATCCGAGCCACATACT 2700
Db 1978 ----- 1977
Qy 2701 CTCGGTGGGCTCCGGTCCCTGGATCACACCCAGGTCCGTGTCGACTACCCGTATAGGC 2760
Db 1978 ----- 1977
Qy 2761 TTTGGCATTATCCTTGTACCATCAACTACACCATATTTAAATCAGGATGTACGTGGAG 2820
Db 1978 ----- 1977
Qy 2821 GGTCGAACACAGGCTGGAAGTGCCTGCAACTGGACGCGGGGGAACGTTGCGATCTGG 2880
Db 1978 ----- 1977
Qy 2881 AAGATAGGNACAGGTCGAGATCGATATGGAGAACATCATCAGGATTCCTAGGACCCC 2940
Db 1978 ----- 1977
Qy 2941 TGCTCGTGTTCAGCGGGGGTTTTTCTTGTGTGACAAGATTCCTCAATACCGCAGATC 3000
Db 1978 ----- 1977

Qy 3001 TAGACTCGTGTGGAGCTTCTCTCAATTTTCTAGGGGGATCTCCCGTGTGTCTTTGGCCAAA 3060
Db 1978 ----- 1977
Qy 3061 AITCGCAGTCCCCAACCTCCAATCACTACCAACCTCCTGTCTCTCAATTTTGTCTGGTT 3120
Db 1978 ----- 1977
Qy 3121 ATCGCTGGATGTCTGTGCGCGTTTTTATCATATTCCTTCATTCCTGTCTATATGCTCA 3180
Db 1978 ----- 1977
Qy 3181 TCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTGGCCGTTTGTCTCTTAATTTCCAG 3240
Db 1978 ----- 1977
Qy 3241 GATCAACAACAACCACTACGGGACCATGCAAAACCTGCACGACTCCTGTCTCAAGGCAACT 3300
Db 1978 ----- 1977
Qy 3301 CTATGTTTCCCTCATGTTGTGTACAAAACCTACCGATGGAATTCACCTGTATTTCCCA 3360
Db 1978 ----- 1977
Qy 3361 TCCCATCGTCTGGGCTTTCGCAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT 3420
Db 1978 ----- 1977
Qy 3421 GGCTCAGTTTACTAGTGCCATTGTTTCAGTGGTTCTAGGGCTTTCGCCACTGTTTGGC 3480
Db 1978 ----- 1977
Qy 3481 TTTACGCTATATGGATGATGTGTATTTGGGGGCCAAGTCTGTACAGCATCGTGAAGTCCCT 3540
Db 1978 ----- 1977
Qy 3541 TTATACCGGTGTACCAATTTTCTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
Db 1978 -----CTAAGAAATTCAGACTCG 1994
Qy 3601 AGCAAGCTAGAAAAGCGCGCCAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA 3660
Db 1995 AGCAAGCTAGAAAAGCGCGCCAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA 2054
Qy 3661 TCAGCCTCGACTGTGCTTCTAGTTGCCAGCATCTGTGTTTGGCCCTTCCCTCCCGTGCCT 3720
Db 2055 TCAGCCTCGACTGTGCTTCTAGTTGCCAGCATCTGTGTTTGGCCCTTCCCTCCCGTGCCT 2114
Qy 3721 TCCTTGACCTGGAAAGGTGCCACTCCCACTGTCTTTCTTAAATGAGGAATTTGCA 3780
Db 2115 TCCTTGACCTGGAAAGGTGCCACTCCCACTGTCTTTCTTAAATGAGGAATTTGCA 2174
Qy 3781 TCGCATTTCTGAGTAGTGTCAATCTATTCTGGGGGGTGGGTGGGCGAGCAGCAAG 3840
Db 2175 TCGCATTTCTGAGTAGTGTCAATCTATTCTGGGGGGTGGGTGGGCGAGCAGCAAG 2234
Qy 3841 GGGGAGATTGGGAAGACAATAGCAGGCATGTGGGAGCTTTCCGCTTCTCGCTCAC 3900
Db 2235 GGGGAGATTGGGAAGACAATAGCAGGCATGTGGGAGCTTTCCGCTTCTCGCTCAC 2294
Qy 3901 TGACTCGCTCGCTCGGTGTTCCGCTGCGGAGCGGTATCAGTCACTCAAGGCGGT 3960
Db 2295 TGACTCGCTCGCTCGGTGTTCCGCTGCGGAGCGGTATCAGTCACTCAAGGCGGT 2354
Qy 3961 AATACGTTTATCCACAGAATCAGGGGATAACGAGGAAAGAAATGTGAGCAAAAGGCCA 4020
Db 2355 AATACGTTTATCCACAGAATCAGGGGATAACGAGGAAAGAAATGTGAGCAAAAGGCCA 2414
Qy 4021 GCAAAAGGCCAGGAACCGTAAAAAGCCGTTGTGGCGTTTTTCCATAGGCTCCGCC 4080
Db 2415 GCAAAAGGCCAGGAACCGTAAAAAGCCGTTGTGGCGTTTTTTCATAGGCTCCGCC 2474

4081 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGAAACCCGACGACT 4140
Db |||||
2475 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGAAACCCGACGACT 2534
Qy |||||
4141 ATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCT 4200
Db |||||
2535 ATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCT 2594
Qy |||||
4201 GCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGAAGCGTGCGCTTCTCAATG 4260
Db |||||
2595 GCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGAAGCGTGCGCTTCTCAATG 2654
Qy |||||
4261 CTCACGCTGAGTATCTCAGTTCGGTGTAGTGTGCTTCCAGCTCCAGCTGGCTGTGCA 4320
Db |||||
2655 CTCACGCTGAGTATCTCAGTTCGGTGTAGTGTGCTTCCAGCTCCAGCTGGCTGTGCA 2714
Qy |||||
4321 CGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAATATCGTCTTGAGTCCA 4380
Db |||||
2715 CGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGTTAATATCGTCTTGAGTCCA 2774
Qy |||||
4381 CCGGTAAGACACGACTTATCGCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGC 4440
Db |||||
2775 CCGGTAAGACACGACTTATCGCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGC 2834
Qy |||||
4441 GAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACACTACGGCTACACTAG 4500
Db |||||
2835 GAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACACTACGGCTACACTAG 2894
Qy |||||
4501 AAGGACAGTATTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGAAGAGTTGG 4560
Db |||||
2895 AAGGACAGTATTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGAAGAGTTGG 2954
Qy |||||
4561 TAGCTCTGATCCGCAAAACAAACCAACCGCTGTGTAGCGTGGTTTTTTTGGTTGCAAGCA 4620
Db |||||
2955 TAGCTCTGATCCGCAAAACAAACCAACCGCTGTGTAGCGTGGTTTTTTTGGTTGCAAGCA 3014
Qy |||||
4621 GCAGATTACCGCGAGAAAAAGGATCTCAAGAAGATCTCTTTGATCTTTTACCGGGTC 4680
Db |||||
3015 GCAGATTACCGCGAGAAAAAGGATCTCAAGAAGATCTCTTTGATCTTTTACCGGGTC 3074
Qy |||||
4681 TGACGCTAGTGGAGCAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAG 4740
Db |||||
3075 TGACGCTAGTGGAGCAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAG 3134
Qy |||||
4741 GATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 4800
Db |||||
3135 GATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 3194
Qy |||||
4801 TGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGAT 4860
Db |||||
3195 TGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGAT 3254
Qy |||||
4861 CTGCTATTTTGGTTCATCATAGTTGCTGACTCTCCCGTGTGTAGATAAATCAGATACG 4920
Db |||||
3255 CTGCTATTTTGGTTCATCATAGTTGCTGACTCTCCCGTGTGTAGATAAATCAGATACG 3314
Qy |||||
4921 GGAGGCTTACCATCTGGCCGACGCTGCAATGATACCGCAGACCCACGCTCACCGG 4980
Db |||||
3315 GGAGGCTTACCATCTGGCCGACGCTGCAATGATACCGCAGACCCACGCTCACCGG 3374
Qy |||||
4981 TCCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGCGCGAGAGTGGTCTCTCG 5040
Db |||||
3375 TCCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGCGCGAGAGTGGTCTCTCG 3434
Qy |||||
5041 AACTTTATCCGCTCCATCAGTCTATTAATTTGTCGCGGAAGCTAGAGTAGTATTC 5100
Db |||||
3435 AACTTTATCCGCTCCATCAGTCTATTAATTTGTCGCGGAAGCTAGAGTAGTATTC 3494
Qy |||||
5101 GCCAGTTATAGTTTGGCAAGCTTGGCTTGCATTTGCTACAGGCATGCTGTCAGCTC 5160
Db |||||
3495 GCCAGTTATAGTTTGGCAAGCTTGGCTTGCATTTGCTACAGGCATGCTGTCAGCTC 3554
Qy |||||
5161 GTCGTTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATC 5220

Db |||||
3555 GTCTTGGTATGGCTTCATTTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATGATC 3614
Qy |||||
5221 CCCCATGTTGTCGAAAAAGCGTTAGCTTCTCGTCTCCGATCGTTGTCGAGAGTAA 5280
Db |||||
3615 CCCCATGTTGTCGAAAAAGCGTTAGCTTCTCGTCTCCGATCGTTGTCGAGAGTAA 3674
Qy |||||
5281 GTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTGAT 5340
Db |||||
3675 GTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTGAT 3734
Qy |||||
5341 GCCATCCGTAAGATGTTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 5400
Db |||||
3735 GCCATCCGTAAGATGTTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAATA 3794
Qy |||||
5401 GTGTATGCGCGACCGAGTTGCTTTCGCGCGGTCAATACGGGATAATACCGCGCCACA 5460
Db |||||
3795 GTGTATGCGCGACCGAGTTGCTTTCGCGCGGTCAATACGGGATAATACCGCGCCACA 3854
Qy |||||
5461 TAGCAGAACTTTAAAGTGTCTCATTCATTGGAACCGTTCTTCGGGCGGAAAACTCTCAAG 5520
Db |||||
3855 TAGCAGAACTTTAAAGTGTCTCATTCATTGGAACCGTTCTTCGGGCGGAAAACTCTCAAG 3914
Qy |||||
5521 GATCTTACCGCTGTTGAGATCCAGTTTCGATTAACCCACTCGTGCACCCAACTGTCTTC 5580
Db |||||
3915 GATCTTACCGCTGTTGAGATCCAGTTTCGATTAACCCACTCGTGCACCCAACTGTCTTC 3974
Qy |||||
5581 AGCATCTTTTACTTTCACCGGCTTCTCGGTGAGCAAAAACAGGAGGAAAAATCCCGC 5640
Db |||||
3975 AGCATCTTTTACTTTCACCGGCTTCTCGGTGAGCAAAAACAGGAGGAAAAATCCCGC 4034
Qy |||||
5641 AAAAAAGGAATAAGGCGCACACGGAATGTTGAATCTCATACTCTTCTTTTCAATA 5700
Db |||||
4035 AAAAAAGGAATAAGGCGCACACGGAATGTTGAATCTCATACTCTTCTTTTCAATA 4094
Qy |||||
5701 TTATTGAAGCATTTATCAGGGTTATGTTCTCATAGCGGATACATATTTGAATGTATTTA 5760
Db |||||
4095 TTATTGAAGCATTTATCAGGGTTATGTTCTCATAGCGGATACATATTTGAATGTATTTA 4154
Qy |||||
5761 GAAAAATAAATAAGGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGCTTA 5820
Db |||||
4155 GAAAAATAAATAAGGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGCTTA 4214
Qy |||||
5821 AGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCCCTTTTCG 5880
Db |||||
4215 AGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCCCTTTTCG 4274
Qy |||||
5881 TC 5882
Db |||||
4275 TC 4276

RESULT 7

AX164583
LOCUS Sequence 5 from Patent WO0138360. 4282 bp DNA linear PAT 22-JUN-2001
DEFINITION
ACCESSION AX164583
VERSION AX164583.1 GI:14545517
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Coit, D.C., Medina-Selby, A.C., Selby, M.C. and Houghton, M.C.
TITLE Novel hcv non-structural polypeptide
JOURNAL Patent: WO 0138360-A 5 31-MAY-2001;
CHIRON CORPORATION (US)
FEATURES
source
1..4282
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCMVII"

ORIGIN

Query Match		44.2%;	Score 2601.2;	DB 6;	Length 4282;		
Best Local Similarity		72.5%;	Pred. No. 0;				
Matches 4268;		Conservative 0;	Mismatches 8;	Indels 1612;	Gaps 5;		
QY	1	TCCGCGTTTCGGTGATGACGGTGAANAACCTCTGACACATGCGCTCCCGGAGACGGTCA	60				
DB	1	TCCGCGTTTCGGTGATGACGGTGAANAACCTCTGACACATGCGCTCCCGGAGACGGTCA	60				
QY	61	CAGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGCGCGGTCTCAGCGGGTG	120				
DB	61	CAGCTTGTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGCGCGGTCTCAGCGGGTG	120				
QY	121	TTGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC	180				
DB	121	TTGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC	180				
QY	181	ACCATATGAAGCTTTTGCAAAAGCCTTAGCCCTCCAAAAGCCTCTCACTACTTCTGG	240				
DB	181	ACCATATGAAGCTTTTGCAAAAGCCTTAGCCCTCCAAAAGCCTCTCACTACTTCTGG	240				
QY	241	AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAAATAAAAAAATTAGTCAGCCA	300				
DB	241	AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAAATAAAAAAATTAGTCAGCCA	300				
QY	301	TGGGCGGAGAA'TGGCGGAACTGGGCGGAGGGAATTATTGGCTATTGGCCATTGCAT	360				
DB	301	TGGGCGGAGAA'TGGCGGAACTGGGCGGAGGGAATTATTGGCTATTGGCCATTGCAT	360				
QY	361	ACGTTGTATCTATATCATATATGTACATTTATTTGGCTATGTCCTCAATATGACGCCA	420				
DB	361	ACGTTGTATCTATATCATATATGTACATTTATTTGGCTATGTCCTCAATATGACGCCA	420				
QY	421	TGTTGACATTGATTATTGACTAGTATTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT	480				
DB	421	TGTTGACATTGATTATTGACTAGTATTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT	480				
QY	481	AGCCCATATATGGAGTTCGCGTTACATAA'CTTACCGTAAATGGCCCGCTGGCTGACCG	540				
DB	481	AGCCCATATATGGAGTTCGCGTTACATAA'CTTACCGTAAATGGCCCGCTGGCTGACCG	540				
QY	541	CCCAAGACCCCGCCCATGAGTCAATATGACGTATGTTCCCATAGTACGCCAATA	600				
DB	541	CCCAAGACCCCGCCCATGAGTCAATATGACGTATGTTCCCATAGTACGCCAATA	600				
QY	601	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	660				
DB	601	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	660				
QY	661	CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGAGCTCAATGACCGTAAATGGCCC	720				
DB	661	CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGAGCTCAATGACCGTAAATGGCCC	720				
QY	721	GCCTGCAATATGCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC	780				
DB	721	GCCTGCAATATGCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC	780				
QY	781	GTATTAGTCAATCGCTATTACATGGTGTATGCGGTTTTGGCAGTACACCAATGGGCGTGA	840				
DB	781	GTATTAGTCAATCGCTATTACATGGTGTATGCGGTTTTGGCAGTACACCAATGGGCGTGA	840				
QY	841	TAGCGGTTTGACTCAGCGGGATTTCAAAGTCTCCACCCCAATGACGTCAATGGGAGTTTG	900				
DB	841	TAGCGGTTTGACTCAGCGGGATTTCAAAGTCTCCACCCCAATGACGTCAATGGGAGTTTG	900				
QY	901	TTTTGGCAACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG	960				
DB	901	TTTTGGCAACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG	960				
QY	961	CAAAATGGGCGGTAGGCGGTACCGGTGGAGGCTCTATATAGCAGAGCTCGTTTAGTGAAC	1020				
DB	961	CAAAATGGGCGGTAGGCGGTACCGGTGGAGGCTCTATATAGCAGAGCTCGTTTAGTGAAC	1020				

QY	1021	CGTCAGATCGCCTGGAGAGCCCATCCACGCTGTTTGTAGCCTCCATAGAAAGACACCCGGAC	1080				
DB	1021	CGTCAGATCGCCTGGAGAGCCCATCCACGCTGTTTGTAGCCTCCATAGAAAGACACCCGGAC	1080				
QY	1081	CGATCCAGCCTCCGGCGCGGAAACGGTGCATTGGAAACGGGATTCGCCGTGCAAGAGT	1140				
DB	1081	CGATCCAGCCTCCGGCGCGGAAACGGTGCATTGGAAACGGGATTCGCCGTGCAAGAGT	1140				
QY	1141	GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCTATGCTATA	1200				
DB	1141	GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCTATGCTATA	1200				
QY	1201	CTGTTTTTGGCTTGGGCGCTATACACCCCGGC-TCCCTATGCTATAGGTATGCTATAGC	1259				
DB	1201	CTGTTTTTGGCTTGGGCGCTATACACCCCGGCTCCCTATGCTATAGGTATGCTATAGC	1260				
QY	1260	TTAGCCTATAGGTGGGTTATTGACCATTTATTGACCACTCCCTATTGGTGACGATACT	1319				
DB	1261	TTAGCCTATAGGTGGGTTATTGACCATTTATTGACCACTCCCTATTGGTGACGATACT	1320				
QY	1320	TTCCATTACTAATCCATAACATGGCTCTTTGGCCACAACCTATCTTATTTGGCTATATGCCA	1379				
DB	1321	TTCCATTACTAATCCATAACATGGCTCTTTGGCCACAACCTATCTTATTTGGCTATATGCCA	1380				
QY	1380	ATACTCTGCTCTCAGAGACTGACAGGACTCTGTATTTTACAGGATGGGT-CCATTT	1438				
DB	1381	ATACTCTGCTCTCAGAGACTGACAGGACTCTGTATTTTACAGGATGGGTCCCATTT	1440				
QY	1439	ATTATTACAATAATTCACATATACAACGCGTCCCGTCCCGTCCCGCAGTCTTTTATATAA	1498				
DB	1441	ATTATTACAATAATTCACATATACAACGCGTCCCGTCCCGTCCCGCAGTCTTTTATATAA	1500				
QY	1499	CATAGCGTGGGATCTCC---GACATCTCGGGTACGTGTTCCGGACATGGCTCTTCTCCG	1555				
DB	1501	CATAGCGTGGGATCTCCACGCGGAACTCTCGGGTACGTGTTCCGGACATGGCTCTTCTCCG	1560				
QY	1556	GTAGCGCGGAGCTTCCACATCGGAGCCCTGGTCCCATCCGTCAGCGGCTCATGTGCG	1615				
DB	1561	GTAGCGCGGAGCTTCCACATCGGAGCCCTGGTCCCATCCGTCAGCGGCTCATGTGCG	1620				
QY	1616	TCGGCAGCTCCTTGTCTTAAACAGTGGAGGCGAGACTTAGGCACAGCAAAATGCCACCA	1675				
DB	1621	TCGGCAGCTCCTTGTCTTAAACAGTGGAGGCGAGACTTAGGCACAGCAAAATGCCACCA	1680				
QY	1676	CCACCAAGTGTCCGCAACAGGCGGTGGCGGTAGGGTATGTCTGAAAATGAGCTCGGAG	1735				
DB	1681	CCACCAAGTGTCCGCAACAGGCGGTGGCGGTAGGGTATGTCTGAAAATGAGCTCGGAG	1740				
QY	1736	ATTGGGCTCGACCC-TGGACGCGAGATGGAAGACTTAAGCGAGCGGCGAGAGATGCAG	1794				
DB	1741	ATTGGGCTCGACCCGCTGACGAGATGGAAGACTTTAAGCGAGCGGCGAGAGATGCAG	1800				
QY	1795	GCAGCTGAGTGTGTGTTATTCTGATAAGAGTCAAGAGTAACTCCCGTTGCGGTGCTGTAA	1854				
DB	1801	GCAGCTGAGTGTGTGTTATTCTGATAAGAGTCAAGAGTAACTCCCGTTGCGGTGCTGTAA	1860				
QY	1855	CGGTGAGGCGCAGTGTAGTCTGAGCAGTACTCGTTGCTCCGCGCGCGCCACACAGACATA	1914				
DB	1861	CGGTGAGGCGCAGTGTAGTCTGAGCAGTACTCGTTGCTCCGCGCGCGCCACACAGACATA	1920				
QY	1915	ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGCTCTTTTCTGACGTACCGTCCGTC	1974				
DB	1921	ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGCTCTTTTCTGACGTACCGTCCGTC	1980				
QY	1975	GACGAATTCAAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGT	2034				
DB	1981	GAC-----	1983				
QY	2035	GTGGAGAGTCTTTCGTTTCCCGCAGCGCTAGCGAAACCCACGCTCACCGGGGGAAGTGC	2094				
DB	1984	-----	1983				

Qy	2095	GCCACACTGTCTGGATTGTGTAGCCTCTCGCACCGGGCCAAACGATCCAGC	2154
Db	1984	-----	1983
Qy	2155	TGATCAACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACTGCAATGATAGCC	2214
Db	1984	-----	1983
Qy	2215	TCAACACCGCTGTGTGGCAGGCTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTC	2274
Db	1984	-----	1983
Qy	2275	CTGAGAGCTAGCCAGCTGCGGACCCCTTACCAGATTTTGAACAGGGCTGGGGCCCTATCA	2334
Db	1984	-----	1983
Qy	2335	GTTATGCCAAGGAAGCGGCCCGACCGCCCTACTGCTGSACTACCCCCCAAAC	2394
Db	1984	-----	1983
Qy	2395	CTTGCGTATTGTGCCCGGAAGAGTGTGTGTCCGTATATTGTTCACTCCAGCC	2454
Db	1984	-----	1983
Qy	2455	CCGTGTGTGGGAACGACCGACAGGTGCGGCGGCCCACTCAGCTGGGTGAAAAATG	2514
Db	1984	-----	1983
Qy	2515	ATAGGACGCTTCGTCTTAACAATACAGGCACCGCTGGGCAATTGGTTCGGTTGTA	2574
Db	1984	-----	1983
Qy	2575	CTTGTGATAACTCAACTGGATTACCAAAGTGTGCGAGCGCTCTTGTGCATCGGAG	2634
Db	1984	-----	1983
Qy	2635	GGCGGGCAACAACACCTGCACTGCCCACTGATTGCTTCGCAAGCATCCGGACGGCCA	2694
Db	1984	-----	1983
Qy	2695	CATACTCTCGTGGGTCCGGTCCCTGGATCACACCCAGGTGCTGTCGACTACCCGT	2754
Db	1984	-----	1983
Qy	2755	ATAGGCTTTGGCAATTATCTTGTACCATCAACTACACATATTAAAAATCAGGATGTACG	2814
Db	1984	-----	1983
Qy	2815	TGGAGGGTGGAAACACAGGCTGGAGTGCCTGCAACTGGACGGGGCGAAGTTGGG	2874
Db	1984	-----	1983
Qy	2875	ATCTGGAAGATAGGACAGGTCCGAGATCGATGGAGAATCACAATCAGGATTCCTAG	2934
Db	1984	-----	1983
Qy	2935	GACCCCTGCTGTGTACAGCGGGTTTTTCTTGTGACAAGAACTCTCACAATACCGC	2994
Db	1984	-----	1983
Qy	2995	AGAGTCTAGACTCGTGTGGACTTCTCTCAATTTTCTAGGGGGATCTCCCGTGTGTTTG	3054
Db	1984	-----	1983
Qy	3055	GCCMAAATTCGAGTCCCAACCTCCAATCACTACCAACCTCTGCTGCTCCAATTTGTC	3114
Db	1984	-----	1983
Qy	3115	CTGTTATCGCTGGATGTGTGCGGGTTTTATCATATTCTCTTCATCTCTGCTGCTAT	3174
Db	1984	-----	1983
Qy	3175	GCCTCATCTTTATTGTTCTCTGGATTATCAAGTATGTTGCCGGTTGTTCCTCTAA	3234

Db	1984	-----	1983
Qy	3235	TTCCAGGATCAACAACCAACAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAG	3294
Db	1984	-----	1983
Qy	3295	GCAACTCTATGTTTCCCTCATGTTGCTGTGTACAAAACCTACGGATGGAAATTGCACCTGTA	3354
Db	1984	-----	1983
Qy	3355	TTCCATCCCATCGTCTCGGCTTTTCGAAAATACCTATGGAGTGGGCTCAGTCCGTT	3414
Db	1984	-----	1983
Qy	3415	TCTCTGGCTCAGTTTACTAGTGCCATTGTTCAGTGTGTCGTAGGGCTTTCCCCCACTG	3474
Db	1984	-----	1983
Qy	3475	TTTGGCTTTCAGCTATATGATGATGTGGTA'TTGGGGGCCAAGTCTGTACAGCATCGTGA	3534
Db	1984	-----	1983
Qy	3535	GTCCCTTTATACCGCTGTACCAATTTCTTTTGTCTCTGGGTATACATTTAAGAAATTC	3594
Db	1984	-----	1984
Qy	3595	GACTCGAGCAAGTCTAGAAAGCGCGCAAGATATCAAGATCCATTTAAGAAATTC	3654
Db	1995	GACTCGAGCAAGTCTAGAAAGCGCGCAAGATATCAAGATCCATTTAAGAAATTC	2054
Qy	3655	CGCTGATCAGCTCGAGTGCCTTTCTAGTTGCCAGCCATCTGTGTTTGCCTCCCTCC	3714
Db	2055	CGCTGATCAGCTCGAGTGCCTTTCTAGTTGCCAGCCATCTGTGTTTGCCTCCCTCC	2114
Qy	3715	GTGCTTCTTGACCTCGAAGGTGCCATCCCACTGTCTCTTCTTAATAAAATGAGGAA	3774
Db	2115	GTGCTTCTTGACCTCGAAGGTGCCATCCCACTGTCTCTTCTTAATAAAATGAGGAA	2174
Qy	3775	ATTGCATCGCATTTGTCTGAGTAGTGTCA'TTCTATCTGGGGGTGGGTGGGCGAGGAC	3834
Db	2175	ATTGCATCGCATTTGTCTGAGTAGTGTCA'TTCTATCTGGGGGTGGGTGGGCGAGGAC	2234
Qy	3835	AGCAAGGGGAGGATTGGGAAGACAATPAGCAGCATGCTCGGGAGCTCTTCCGCTTCTC	3894
Db	2235	AGCAAGGGGAGGATTGGGAAGACAATPAGCAGCATGCTCGGGAGCTCTTCCGCTTCTC	2294
Qy	3895	GCTCACTGACTCGCTCGCTCGGTTCGGCTTCGGCTCGCGCGAGCGGTATCAGCTCACTCAA	3954
Db	2295	GCTCACTGACTCGCTCGCTCGGTTCGGCTTCGGCTCGCGCGAGCGGTATCAGCTCACTCAA	2354
Qy	3955	GGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGAGGAAGACATGTGAGCAAA	4014
Db	2355	GGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGAGGAAGACATGTGAGCAAA	2414
Qy	4015	AGGCCAGAAAAGSCCAGGAACCGTAAAGGCGCGTTCCTGGCGTTTTTCCATAGGCT	4074
Db	2415	AGGCCAGAAAAGSCCAGGAACCGTAAAGGCGCGTTCCTGGCGTTTTTCCATAGGCT	2474
Qy	4075	CCGCCCCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGAC	4134
Db	2475	CCGCCCCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGAC	2534
Qy	4135	AGGACTAATAGATACCAGAGCTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCC	4194
Db	2535	AGGACTAATAGATACCAGAGCTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCC	2594
Qy	4195	GACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAGCGTGGCGCTTC	4254
Db	2595	GACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAGCGTGGCGCTTC	2654
Qy	4255	TCAATGCTCAAGCTGTAGGTATCTCAGTTGGTGTAGGTCGTTCCAAAGCTGGGCTG	4314

```
Db      2655  TCAATGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGCGTG 2714
QY      4315  TGTGCAAGAACCCCGGTTTCAAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGA 4374
Db      2715  TGTGCAAGAACCCCGGTTTCAAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGA 2774
QY      4375  GTCCAAACCGGTAAGACACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAG 4434
Db      2775  GTCCAAACCGGTAAGACACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAG 2834
QY      4435  CAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGTCCTAACTACGGCTA 4494
Db      2835  CAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGTCCTAACTACGGCTA 2894
QY      4495  CACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGAAAAAAG 4554
Db      2895  CACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGAAAAAAG 2954
QY      4555  AGTTGGTAGCTCTTGTATCGGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGTGTTG 4614
Db      2955  AGTTGGTAGCTCTTGTATCGGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTGTGTTG 3014
QY      4615  CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTTTCTAC 4674
Db      3015  CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTCTTGTATCTTTCTAC 3074
QY      4675  GGGGCTCTGACGCTCAGTGGACGAAAAACTCACGTTAAGGATTTTGGTCAATGAGATTATC 4734
Db      3075  GGGGCTCTGACGCTCAGTGGACGAAAAACTCACGTTAAGGATTTTGGTCAATGAGATTATC 3134
QY      4735  AAAAAAGATCTTCACTAGATCTCTTTAAATTAATAAGTAAAGTTTAAATCAATCTAAAG 4794
Db      3135  AAAAAAGATCTTCACTAGATCTCTTTAAATTAATAAGTAAAGTTTAAATCAATCTAAAG 3194
QY      4795  TATATATGAGTAACTTTGGTCTCAGAGTTACCAATCTTAATCAGTGAGGCACCTATCTC 4854
Db      3195  TATATATGAGTAACTTTGGTCTCAGAGTTACCAATCTTAATCAGTGAGGCACCTATCTC 3254
QY      4855  AGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGTCTGTGTAGATAACTAC 4914
Db      3255  AGCGATCTGTCTATTTTCGTTTCACTCCATAGTGTGCTGACTCCCGTCTGTGTAGATAACTAC 3314
QY      4915  GATACGGAGGGCTTACCATCTGCGCCCGCAGTGTGCAATGATACCGGAGACCCAGCTC 4974
Db      3315  GATACGGAGGGCTTACCATCTGCGCCCGCAGTGTGCAATGATACCGGAGACCCAGCTC 3374
QY      4975  ACCGGCTCCAGATTATTCAGCAATAAACACGAGCCAGCGGAGGCGAGCGAGAGTGG 5034
Db      3375  ACCGGCTCCAGATTATTCAGCAATAAACACGAGCCAGCGGAGGCGAGCGAGAGTGG 3434
QY      5035  TCCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAG 5094
Db      3435  TCCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAG 3494
QY      5095  TAGTTCCGCAAGTTAATAGTTTGGCAACGTTTGGCCATGCTACAGGCATCTGTGTGTC 5154
Db      3495  TAGTTCCGCAAGTTAATAGTTTGGCAACGTTTGGCCATGCTACAGGCATCTGTGTGTC 3554
QY      5155  ACCGCTCGTCTTTGGTATGGCTTTCATTTCAGCTCCGGTTCGCAACGATCAAGGCGAGTTAC 5214
Db      3555  ACCGCTCGTCTTTGGTATGGCTTTCATTTCAGCTCCGGTTCGCAACGATCAAGGCGAGTTAC 3614
QY      5215  ATGATCCCCCATGTTGTGCAAAAAAGGGTTAGTCTCTTCGGTCTTCGATCGTTGTCTAG 5274
Db      3615  ATGATCCCCCATGTTGTGCAAAAAAGGGTTAGTCTCTTCGGTCTTCGATCGTTGTCTAG 3674
QY      5275  AAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTAC 5334
Db      3675  AAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTAC 3734
QY      5335  TGTCAAGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG 5394
Db      3735  TGTCAAGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG 3794
```

```
QY      5395  AGAATAGTGTATGCGGCGAACCGAGTTGCTCTTGCCTCCGCGGTCAATACGGGATATACCGC 5454
Db      3795  AGAATAGTGTATGCGGCGAACCGAGTTGCTCTTGCCTCCGCGGTCAATACGGGATATACCGC 3854
QY      5455  GCCACATACGACAACTTTTAAAGATGCTCATCATTTGGAACACGTTCTTCGGGGGGAACACT 5514
Db      3855  GCCACATACGACAACTTTTAAAGATGCTCATCATTTGGAACACGTTCTTCGGGGGGAACACT 3914
QY      5515  CTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCCGATGTAAACCCACTCGTGCAACCACTG 5574
Db      3915  CTCAAGGATCTTTACCGCTGTTGAGATCCAGTTTCCGATGTAAACCCACTCGTGCAACCACTG 3974
QY      5575  ATCTTCAGCATCTTTTACTTTTCCAGCGTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGAAAA 5634
Db      3975  ATCTTCAGCATCTTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGAAAA 4034
QY      5635  TCCCGCAAAAAGGGAATAAGGCGGACACGGAATGTTTCAATACTCATACTCTTCCTTTT 5694
Db      4035  TCCCGCAAAAAGGGAATAAGGCGGACACGGAATGTTTCAATACTCATACTCTTCCTTTT 4094
QY      5695  TCAATATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTCGAATG 5754
Db      4095  TCAATATTATTGAAGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTCGAATG 4154
QY      5755  TATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGGAAAAAGTGCCACTGA 5814
Db      4155  TATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGGAAAAAGTGCCACTGA 4214
QY      5815  CGTCTAAGAAACCATTTATTATCATGACATTAACTATAAAATAGCGGTATCACGAGGCC 5874
Db      4215  CGTCTAAGAAACCATTTATTATCATGACATTAACTATAAAATAGCGGTATCACGAGGCC 4274
QY      5875  CTTTCGTC 5882
Db      4275  CTTTCGTC 4282

RESULT 8
LOCUS   CQ855138                6050 bp    DNA        linear    PAT 23-AUG-2004
DEFINITION   Sequence 4 from Patent WO2004058808.
ACCESSION   CQ855138
VERSION     CQ855138.1   GI:51510567
KEYWORDS    .
SOURCE      synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
            Hooper,J.W., Schmaljohn,C.S. and Custer,M.
            Dna vaccines against hantavirus infections
            Patent: WO 2004058808-A 4 15-JUL-2004;
            U.S. Army Medical Research Institute of Infectious Diseases es
            Department of the Army (US)
FEATURES             Location/Qualifiers
     source          1..6050
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pWRG-SEO-S = Seoul virus S segment, strain SR-11,
                     subcloned into DNA vector pWRG7077"

ORIGIN
Query Match      40.8%; Score 2397.8; DB 6; Length 6050;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 3456; Conservative 0; Mismatches 1097; Indels 240; Gaps 16;

QY      337  ATTATTTGGCTATTTGGCCATTGCACTGTTATCTATATATCATATATGATGACATTATTT 396
Db      1263  AATATTTGGCTATTTGGCCATTGCACTGTTATCTATATATCATATATGATGACATTATTT 1322
QY      397  GGCTCATGTCCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAAATAGTAA 456
            |||||
```

[illegible]

1532	Qy	GTTCGGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCCACATCCGAGCCCTGTGTC	1599	Db	GTTCGGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCCACATCCGAGCCCTGTGTC
2462	Db	GTTCGGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCCACATCCGAGCCCTGTGTC	2521	Qy	ATCCGTCACAGCGCTCATGGTCGCTCGGAGCTCTTGTGCTCTTAAACAGTGGAGGCCAGAC
1592	Qy	ATCCGTCACAGCGCTCATGGTCGCTCGGAGCTCTTGTGCTCTTAAACAGTGGAGGCCAGAC	1651	Db	ATGCCTCCAGCGCTCATGGTCGCTCGGAGCTCTTGTGCTCTTAAACAGTGGAGGCCAGAC
2522	Db	ATGCCTCCAGCGCTCATGGTCGCTCGGAGCTCTTGTGCTCTTAAACAGTGGAGGCCAGAC	2581	Qy	TTTAGGCACAGCAAAATGCCCCACACACACACAGTGTGGCGCAAAAGCGCGTAGGGT
1652	Qy	TTTAGGCACAGCAAAATGCCCCACACACACAGTGTGGCGCAAAAGCGCGTAGGGT	1711	Db	TTAGGCACAGCAAAATGCCCCACACACAGTGTGGCGCAAAAGCGCGTAGGGT
2582	Db	TTAGGCACAGCAAAATGCCCCACACACAGTGTGGCGCAAAAGCGCGTAGGGT	2641	Qy	ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCGAGATGGAAGACTTAA
1712	Qy	ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCGAGATGGAAGACTTAA	1771	Db	ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCGAGATGGAAGACTTAA
2642	Db	ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCGAGATGGAAGACTTAA	2701	Qy	GGCAGCGGCAGAAAGATGACAGCGAGCTGAGTTGTGTATTTCTGATAAGAGTCAGAGGT
1772	Qy	GGCAGCGGCAGAAAGATGACAGCGAGCTGAGTTGTGTATTTCTGATAAGAGTCAGAGGT	1831	Db	GGCAGCGGCAGAAAGATGACAGCGAGCTGAGTTGTGTATTTCTGATAAGAGTCAGAGGT
2702	Db	GGCAGCGGCAGAAAGATGACAGCGAGCTGAGTTGTGTATTTCTGATAAGAGTCAGAGGT	2761	Qy	AATCTCCGTTGCGGTGCTGTTAAACGTTGGAGGCGAGTGTAGTCTGTGACGAGTACTCGTTGC
1832	Qy	AATCTCCGTTGCGGTGCTGTTAAACGTTGGAGGCGAGTGTAGTCTGTGACGAGTACTCGTTGC	1891	Db	AATCTCCGTTGCGGTGCTGTTAAACGTTGGAGGCGAGTGTAGTCTGTGACGAGTACTCGTTGC
2762	Db	AATCTCCGTTGCGGTGCTGTTAAACGTTGGAGGCGAGTGTAGTCTGTGACGAGTACTCGTTGC	2821	Qy	TGCGGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTTCCTTTCCATGGG
1892	Qy	TGCGGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTTCCTTTCCATGGG	1951	Db	TGCGGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTTCCTTTCCATGGG
2822	Db	TGCGGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTTCCTTTCCATGGG	2881	Qy	GCTCTGCTGTGCTGCTGCTGTGTGAGGAGCT--TCGTTTTCGCCACGCGCTAGCGA
1952	Qy	GCTCTGCTGTGCTGCTGCTGTGTGAGGAGCT--TCGTTTTCGCCACGCGCTAGCGA	2011	Db	TCTTTTCTGAGTACACGCTGTCGATCACTAACAGAAATGGCAACTATGAGAAATCCAGAGA
2882	Db	TCTTTTCTGAGTACACGCTGTCGATCACTAACAGAAATGGCAACTATGAGAAATCCAGAGA	2937	Qy	GAATCAGTGCTCAGCAGGGGCGAGCTGTGATAGCAGCGCCAGAAAGTCAAGGATCGAGAA
2012	Qy	GAATCAGTGCTCAGCAGGGGCGAGCTGTGATAGCAGCGCCAGAAAGTCAAGGATCGAGAA	2068	Db	ACCAGGGCCACAGCAGACAGTCCAGCTGATCAACACACAGCGGAGTGGCACCTCAATAG
2938	Db	ACCAGGGCCACAGCAGACAGTCCAGCTGATCAACACACAGCGGAGTGGCACCTCAATAG	2997	Qy	AAGCAGTATGAGAAAGATCTCTGATGACTTAAACAGAGGGGCACTGCATGATCGGGAGAGT
2069	Qy	AAGCAGTATGAGAAAGATCTCTGATGACTTAAACAGAGGGGCACTGCATGATCGGGAGAGT	2128	Db	CAGGGCCCTGAACTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCA
2998	Db	CAGGGCCCTGAACTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCA	3057	Qy	GTCCGAGCTTCAATCAATCAAAATTTGATGATTTGAAGGCCAACTTGCCGACAGATTG
2129	Qy	GTCCGAGCTTCAATCAATCAAAATTTGATGATTTGAAGGCCAACTTGCCGACAGATTG	2188	Db	CCACAAGTTCAACTCTTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGA
3058	Db	CCACAAGTTCAACTCTTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGA	3117	Qy	CAGCAGGGAAGACATCCCGGCGAGGACCGGGATCTCTACAGGGGTAGAGCCAGGTGATCAT
2189	Qy	CAGCAGGGAAGACATCCCGGCGAGGACCGGGATCTCTACAGGGGTAGAGCCAGGTGATCAT	2248	Db	TTTGTGACCGGGCTGGGGCCCTATAGTTATGCAACGGAAGCGGCCCGCCAGC---CAGC
3118	Db	TTTGTGACCGGGCTGGGGCCCTATAGTTATGCAACGGAAGCGGCCCGCCAGC---CAGC	3177	Qy	CTTAAGAAAGATCAGCACTAAGCTACGGGAATACACTGGAACCTGATGTCTTGACATT
2249	Qy	CTTAAGAAAGATCAGCACTAAGCTACGGGAATACACTGGAACCTGATGTCTTGACATT	2308	Db	GCCCTACTGCTGGCACTACCCCCAAAACTTCGCGTATTTGTGCCCGGAAGAGTGTGT
3178	Db	GCCCTACTGCTGGCACTACCCCCAAAACTTCGCGTATTTGTGCCCGGAAGAGTGTGT	3237	Qy	GATGAACCTACAGGACAGACAGCTGATTTGGCTGACCAATAATTGCTCTATCTGACATCATTC
2309	Qy	GATGAACCTACAGGACAGACAGCTGATTTGGCTGACCAATAATTGCTCTATCTGACATCATTC	2364	Db	GTGGTCCGG-----TATATTGCTTCACTCCAGCGCCCGCTGGTGGTGGGA
3238	Db	GTGGTCCGG-----TATATTGCTTCACTCCAGCGCCCGCTGGTGGTGGGA	2424	Qy	ACGACCGCAGAGTTCGGGCGCGCCCACTACAGCTGGGGTGAATAATGATACGGACGCTTTC
2365	Qy	ACGACCGCAGAGTTCGGGCGCGCCCACTACAGCTGGGGTGAATAATGATACGGACGCTTTC	2468	Db	GAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG
3298	Db	GAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG	3357	Qy	GTCCCTTAAACATACAGGCGCACCGCTGGGGCAATTGGTTGGTTGATACCTGGATGAACCTCA
2425	Qy	GTCCCTTAAACATACAGGCGCACCGCTGGGGCAATTGGTTGGTTGATACCTGGATGAACCTCA	2468	Db	GGAAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG
3358	Db	GGAAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG	3417	Qy	ACGACCGCAGAGTTCGGGCGCGCCCACTACAGCTGGGGTGAATAATGATACGGACGCTTTC
2469	Qy	ACGACCGCAGAGTTCGGGCGCGCCCACTACAGCTGGGGTGAATAATGATACGGACGCTTTC	2528	Db	AAGACAACAAAGGGATGAGGATCAGATTTCAAGATGACAGCTCATATGAGGATGTCAAT
3418	Db	AAGACAACAAAGGGATGAGGATCAGATTTCAAGATGACAGCTCATATGAGGATGTCAAT	2588	Qy	GTCCCTTAAACATACAGGCGCACCGCTGGGGCAATTGGTTGGTTGATACCTGGATGAACCTCA
2529	Qy	GTCCCTTAAACATACAGGCGCACCGCTGGGGCAATTGGTTGGTTGATACCTGGATGAACCTCA	2598	Db	GGAAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG
3478	Db	GGAAATCAGAAAGCCCAACATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG	3537	Qy	

Db	5698	CCTTCGGAAAAAGAGTGTGTTAGCTCTTGATCCCGCAAAACAAAACACCACGCTGGTAGCCGGTG	5755
Qy	4602	GTTTTTTTGTGTTGAAGCAGCAGATTACGGCGCAGAAAAAAGGATCTCAAGAAGATCCITT	4661
Db	5758	GTTTTTTTGTGTTGAAGCAGCAGATTACGGCGCAGAAAAAAGGATCTCAAGAAGATCCITT	5817
Qy	4662	TGATCTTTTCTACGGGTCTGACCTCAGTGGGAACGAAAACTACGTTAAGGGATTTTGG	4721
Db	5818	TGATCTTTTCTACGGGTCTGACCTCAGTGGGAACGAAAACTACGTTAAGGGATTTTGG	5877
Qy	4722	TCATGAGATTATCAAAAAGGATCTTACCTAGATGCCCTTTTAAATTAATAAATGAAGTTTTA	4781
Db	5878	TCATGAGATTATCAAAAAGGATCTTACCTAGATGCCCTTTTAAATTAATAAATGAAGTTTTA	5937
Qy	4782	AATCAATCTAAAGTATATATAGTAAACTTTGGTCTGCAGCTTACCAATGCTTAATCAGTG	4841
Db	5938	AATCAATCTAAAGTATATATAGTAAACTTTGGTCTGCAGCTTACCAATGCTTAATCAGTG	5997
Qy	4842	AGGCACCTATCTCAGCATCTGCTCTATTGTTTCATCCATAGTTGCCTGACT	4894
Db	5998	AGGCACCTATCTCAGCATCTGCTCTATTGTTTCATCCATAGTTGCCTGACT	6050
RESULT 9			
CQ795336			
LOCUS			
DEFINITION			
Sequence 14 from Patent WO2004024904.			
ACCESSION			
CQ795336			
VERSION			
CQ795336.1 GI:46407431			
KEYWORDS			
synthetic construct			
SOURCE			
ORGANISM			
synthetic construct			
other sequences; artificial sequences.			
REFERENCE			
1			
Bartsch,B. and Cosset,F.L.			
Infectious hepatitis pseudo-particles containing functional el.e2			
envelope proteins			
AUTHORS			
TITLE			
Patent: WO 2004024904-A 14 25-MAR-2004;			
JOURNAL			
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM)			
(FR)			
FEATURES			
source			
Location/Qualifiers			
1..7551			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="phCMV-DCEI2p7 plasmid"			
ORIGIN			
Query Match			
Best Local Similarity 39.7%; Score 2333.2; DB 6; Length 7551;			
Matches 3122; Conservative 0; Mismatches 683; Indels 259; Gaps 12;			
Qy	2018	CTGTGCTCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGCGCTAGCGAAACCCACGT	2077
Db	2832	CTGGCGAAAGTCTCTGTAGTGTCTGTTATTATTCGCCGCGCTCGACGCGAAACCCACGT	2891
Qy	2078	CACCGGGGAAGTCCCGGCCACACTGTCTGTGAATTTGTAGCTCTCTCCACACAGCGCG	2137
Db	2892	CACCGGGGAAGTCCCGGCCCACACCGCTGTGGCTTTGTGTCTCTTACACAGCGCG	2951
Qy	2138	CAAGCAGAACGTCCAGCTGATCAACACCAACGGCAGTTGGCACTCAATAGCAGCGCCCT	2197
Db	2952	CAAGCAGAACATCCAACTGATCAACACCAACGGCAGTTGGCACATCAATAGCAGCGCCT	3011
Qy	2198	GAATGCAATGATAGCTTCAACACCGCTGTGGCAGGGCTTTTCTATCACCACAAGTT	2257
Db	3012	GAATGCAACGATAGCTTCAACACCGCTGTGGCAGGGCTCTCTCTATCGCCACAAT	3071
Qy	2258	CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTTGACCA	2317
Db	3072	CAACTCTTCAGGCTGTCTGAGAGGTTGGCAGCTGCCGAGCTTACCGATTTTGCCCA	3131
Qy	2318	GGGCTGGGGCCCTATCATGTTATGCAACGAAGCGGCCCGACGAGCGCCCTACTGCTG	2377

[illegible]

Db 4212 TGCGTGGTGGCAATGCCCTGGCTCAACAAATACCAGTGAATCTTTTCCCTCTGCCA 4271
Qy 3288 -----GCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAACCTAC 3334
Db 4272 AAAATTATGGGACATCATGAAGCCCTTGAGCATCTGACTTCTGGCTAATAAGGAAAT 4331
Qy 3335 GGATGGAATATGACCTGTATCCCATCCATCGTCTGGGCTTTCGCAAAATACCTATG 3394
Db 4332 TTATTTTCATTGCAATATGATGTTGGAAATTTTGTGCTCTCACTCGGAAGACATATG 4391
Qy 3395 GGAGTGGGCTCAGTCGGTTTCTTGGCTCAGTTTACTAGTGCCATTTGTTCCAGTGGTT 3454
Db 4392 GGAGGCAATCATTTAAACATCAGAAATGAGTATTTGGTTAGAGTTTGGCAACATATG 4451
Qy 3455 CGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGCTATATGAGT- -----ATG 3500
Db 4452 CCCATATGCTGGCTGCCAACAAGGTTGGCTATAAAGAGGTCATCAGTATATGAAC 4511
Qy 3501 TGGTATTTGGGGCCAAAGTCTGTACAGATCGTGAGTCCCTTATACGGCTGTACCAAT 3560
Db 4512 AGCCCCCTGCTGCCATTCCTTATCCATAGAAAGCCTTGACTTGAGGTTAGATTTTTT 4571
Qy 3561 TTCTTTTGTCTCGGTATACATTT-----TAAGAAATTCAGACTCGAGCAA 3605
Db 4572 TTATATTTTGTGTTGTTATTTTTTCTTTAAATCCCTAAATTTTCTTACATGTTT 4631
Qy 3606 GTCTAGAAAGGCGCCAAAGATATCAAGGATCCACTACGGTTAGAGCTCGCTGATC--- 3662
Db 4632 TACTAGCCAGATTTTTCCTCTCTGCTACTACTCCAGTCATAGCTGCCCTCTCTCT 4691
Qy 3663 -----AGCTCGACTGTCCTTCTAGTTGCCAGCC 3692
Db 4692 TATGAGATCCCTCGACGATCGCGCAATTCGTAATCATGTCATAGCTGTTCTCTGTG 4751
Qy 3693 ATCTGTTGTTGGCCCTCCCGCTGCTCTCTTGACCCCTGGAAGTGCCACTCCACTGT 3752
Db 4752 TGAATTTGTTATCCGTTCACAATTCACACAATAGAGCCGGAAGCATAAAGTGTA 4811
Qy 3753 CTTTCTCTAATAAATGAGGAAATTCATGCGCATTTGTTCTGAGTGTGTCATTCTATTCT 3812
Db 4812 GCCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCGCT 4871
Qy 3813 GGGGGTGGGTG---GGCAGACAGCAAGGGGGAGGATTGGGAAGACATAGCAGGCA 3869
Db 4872 TTCAGTTCGGGAAACCTGTGTCGCCAGCTGCATTAAATGAATCGCCCAACGCGCGGAGA 4931
Qy 3870 TGC-----TGGGAGCTCTTCGGCTTCTCGCTCACTCACTCGCTCGCTCGGTG 3919
Db 4932 GGGGTTTGGGTATTGGGCGCTCTTCGGCTTCTCGCTCACTCACTCGCTCGCTCGGTG 4991
Qy 3920 GTTCGGCTGGCGAGCGGTATCAGCTCACTCAAGCGGTATAACGGTTATCCACAGAA 3979
Db 4992 GTTCGGCTGGCGAGCGGTATCAGCTCACTCAAGCGGTATAACGGTTATCCACAGAA 5051
Qy 3980 TCAGGGATACACAGGAAGACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGT 4039
Db 5052 TCAGGGATACACAGGAAGACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGT 5111
Qy 4040 AAAAGGCCGCTGTCTGGGTTTTTCCATAGGCTCCGCCCTCGCTGACAGCATCACAA 4099
Db 5112 AAAAGGCCGCTGTCTGGGTTTTTCCATAGGCTCCGCCCTCGCTGACGAGCATCACAA 5171
Qy 4100 AATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCGAGGTTT 4159
Db 5172 AATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACCGAGGTTT 5231
Qy 4160 CCGCTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTTACCGATACCTG 4219
Db 5232 CCGCTGGAGCTCCCTCGTGGCTCTCTGTTCCGACCTCGCTTACCGATACCTG 5291
Qy 4220 TCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTC 4279
Db 5292 TCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC 5351

Qy 4280 AGTTCCGTTAGTTCGTTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGGTTTCAAGCC 4339
Db 5352 AGTTCCGTTAGTTCGTTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGGTTTCAAGCC 5411
Qy 4340 GACCGCTGGCGCTTATCCGTTAACTATCTCTTGTAGTCCAAACCGGTAGACACAGCTTA 4399
Db 5412 GACCGCTGGCGCTTATCCGTTAACTATCTCTTGTAGTCCAAACCGGTAGACACAGCTTA 5471
Qy 4400 TCGCCACTGGCAGCAGCCTGTTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGCT 4459
Db 5472 TCGCCACTGGCAGCAGCCTGTTAAACAGATTAGCAGAGCGAGGTATGTAGGCGGTGCT 5531
Qy 4460 ACAGATTTCTGAAGTGGTGGCTTAACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCAAA 4519
Db 5532 ACAGATTTCTGAAGTGGTGGCTTAACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCAAA 5591
Qy 4520 TCGGCTCTCTGAAGCAGTACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCAAA 4579
Db 5592 TCGGCTCTCTGAAGCAGTACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCAAA 5651
Qy 4580 CAAACCCAGCTGTAGCGGTGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAA 4639
Db 5652 CAAACCCAGCTGTAGCGGTGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAGAAAA 5711
Qy 4640 AAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGCTGACCGCTCAGTGGAAAGAA 4699
Db 5712 AAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGCTGACCGCTCAGTGGAAAGAA 5771
Qy 4700 AACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCCTT 4759
Db 5772 AACTCAGTTAAGGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCCTT 5831
Qy 4760 TTAATTTAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC 4819
Db 5832 TTAATTTAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC 5891
Qy 4820 AGTTACCAATGCTTAATCAGTAGGACCTATCTAGCGATCTGTCTATTTTGGTTCATCC 4879
Db 5892 AGTTACCAATGCTTAATCAGTAGGACCTATCTAGCGATCTGTCTATTTTGGTTCATCC 5951
Qy 4880 ATAGTTGCTGACTCCCGTCTGTAGATAACTACGATACGGAGGCTTACCATCTGGC 4939
Db 5952 ATAGTTGCTGACTCCCGTCTGTAGATAACTACGATACGGAGGCTTACCATCTGGC 6011
Qy 4940 CCCAGTGTGCAATGATACCGGAGACCCACGCTCAACGGCTCCAGATTATCAGCAATA 4999
Db 6012 CCCAGTGTGCAATGATACCGGAGACCCACGCTCAACGGCTCCAGATTATCAGCAATA 6071
Qy 5000 AACGACGAGCGGAAGGCGCAGCGCAGAGTGGTCTGCAACTTATCCGCTCCATC 5059
Db 6072 AACGACGAGCGGAAGGCGCAGCGCAGAGTGGTCTGCAACTTATCCGCTCCATC 6131
Qy 5060 CAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTTCGCGAGTTAATAGTTTGGCGC 5119
Db 6132 CAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTTCGCGAGTTAATAGTTTGGCGC 6191
Qy 5120 AACGTTGTTGCCATGCTACAGGCATCTGTGGTGTACGCTCGTGGTTGGTATGGGTTCA 5179
Db 6192 AACGTTGTTGCCATGCTACAGGCATCTGTGGTGTACGCTCGTGGTTGGTATGGGTTCA 6251
Qy 5180 TTCAGTCCGGTTCACACGATCAAGCGAGTTACATGATCCCGATGTTGTGCAAAAA 5239
Db 6252 TTCAGTCCGGTTCACACGATCAAGCGAGTTACATGATCCCGATGTTGTGCAAAAA 6311
Qy 5240 GC-GGTTAGTCTCTCGTCCCGATCGTTGTGCAAGTAAGTTGGCGCAGTGTATC 5298
Db 6312 GCGGTTAGTCTCTCGTCCCGATCGTTGTGCAAGTAAGTTGGCGCAGTGTATC 6371
Qy 5299 ACTCATGTTTGGCAGCCTGCATAAATTTCTTACTGTATGCCATCCGTAAAGTGTCTT 5358
Db 6372 ACTCATGTTTGGCAGCCTGCATAAATTTCTTACTGTATGCCATCCGTAAAGTGTCTT 6431

QY 5359 TTCTGTGACTGTGAGTACTCAACCAAGTCAITCTGAGATAGTGTATGCGGACCGAG 5418
DB 6432 TTCTGTGACTGTGAGTACTCAACCAAGTCAITCTGAGATAGTGTATGCGGACCGAG 6491
QY 5419 TTGCTCTTGGCCGGGTCAATACGGGATATACCGGCCACATAGCAGAACTTTAAAGT 5478
DB 6492 TTGCTCTTGGCCGGGTCAATACGGGATATACCGGCCACATAGCAGAACTTTAAAGT 6551
QY 5479 GCTCATCTTGGAAAGCTTTCTCGGGCGAAACCTCTCAAGGATCTTACCGCTGTGTAG 5538
DB 6552 GCTCATCTTGGAAAGCTTTCTCGGGCGAAACCTCTCAAGGATCTTACCGCTGTGTAG 6611
QY 5539 ATCCAGTTCGATGTAAACCCACTCTGTCACCCAACTGATCTTACGATCTTTTACTTTTAC 5598
DB 6612 ATCCAGTTCGATGTAAACCCACTCTGTCACCCAACTGATCTTACGATCTTTTACTTTTAC 6671
QY 5599 CAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGAATAAGGCG 5658
DB 6672 CAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGAATAAGGCG 6731
QY 5659 GACACGGAATGTGAATCTACTACTCTTCTCTTTTCAATATTTTGAAGCATTTATCA 5718
DB 6732 GACACGGAATGTGAATCTACTACTCTTCTCTTTTCAATATTTTGAAGCATTTATCA 6791
QY 5719 GGCTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGG 5778
DB 6792 GGCTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAGG 6851
QY 5779 GGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGAGCTTAAG 5822
DB 6852 GGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTTAAGTGAAG 6895

RESULT 10
CQ867240
LOCUS CQ867240 7551 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 14 from Patent EP1454989.
ACCESSION CQ867240
VERSION CQ867240.1 GI:51997482
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Bartosch, B. and Cosset, F. L.
TITLE Infectious HCV pseudo-particles containing native functional E1 and E2 envelope proteins
JOURNAL Patent: EP 1454989-A 14 08-SEP-2004;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM) (FR)
FEATURES
source Location/Qualifiers
1. 7551
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pHCMV-DCB1E2p7 plasmid"

ORIGIN
Query Match 39.7%; Score 2333.2; DB 6; Length 7551;
Best Local Similarity 76.8%; Pred. No. 0;
Matches 3122; Conservative 0; Mismatches 683; Indels 259; Gaps 12;
QY 2018 CTGTGTGCTGCTGTGTGGAGCAGTCTTCGTTTCCCGCAGCGCTAGCGAAACCCACGT 2077
DB 2832 CTGGGCGAAGGTCTGTGTAGTGTCTGTATTTTGGCGGCTGACGCGGAAACCCACGT 2891
QY 2078 CACCGGGGAAGTCCGGCCACACTGTGTCTGATTTTGTAGGCTCTCTCGCACCGGCGC 2137
DB 2892 CACCGGGGAAGTCCGGCCACACTGACCGCTGGGCTGTGTGTCTCTTACACGAGGCG 2951
QY 2138 CAAGCAGAACGTCAGCTGATCAACCAACAGCGAGTTGGCACTCAATAGCAGCGCCT 2197
DB 2952 CAAGCAGAACATCAACTGATCAACCAACAGCGAGTTGGCACATCAATAGCAGCGCCT 3011

QY 2198 GAACTGCAATGATAGTCTCAACACCGCTGGTTGGCAGGCTTTTCTATCACCACAAGTT 2257
DB 3012 GAACTGCAACGATAGCTTTACCAACCGCTGGTTAGCAGGCTCTTCTATGCGCCACAAT 3071
QY 2258 CAACTCTTTCAGGCTGTCTGAGAGGTACCGAGCTGCCGACCCCTTACCGATTTGACCA 2317
DB 3072 CAACTCTTTCAGGCTGTCTGAGAGGTGGCCAGCTGCCGACCCCTTACCGATTTGCGCCA 3131
QY 2318 GGGCTGGGCGCTATCAGTTATGCCAACGGAAGCGGCCCGCAGCGCCCTTACTGCTG 2377
DB 3132 GGGCTGGGCGCTCATCAGTTATGCCAACGGAAGCGGCCCTTGACGAACGCCCTTACTGTTG 3191
QY 2378 GCACTACCCCCCAAAACCTTTGGCGTATTGTCGCCGGAAGAGTGTGTGGTCCGGTATA 2437
DB 3192 GCACTACCCCTCCAGACCTTTGTGGCATTGTGCCCGCAAGAGCGTGTGTGGCCCGGTATA 3251
QY 2438 TTGCTTCACTCCAGCGCCGTGTGTGGGAACGACGACGAGTCCGGCGCGCCACCTTA 2497
DB 3252 TTGCTTCACTCCAGCGCCGTGTGTGGGAACGACGACGAGTCCGGCGCGCCCTACCTA 3311
QY 2498 CAGCTGGGCGTGAATGATACGAGCTCTTCGTCCTTAACATACCAAGCCACCGCTGG 2557
DB 3312 CAGCTGGGCGTGAATGATACGAGTGTCTTCGTCCTTAACCAACAGCGCCACCGCTGG 3371
QY 2558 CAATTGGTTCGGTGTGATCTGATGAATCAACTGATTTACCAAAAGTGTCCGAGCGCC 2617
DB 3372 CAATTGGTTCGGTGTGATCTGATGAATCAACTGATTTACCAAAAGTGTCCGAGCGCC 3431
QY 2618 TCCTTGTGTATCGGAGGGCGGGCAACAAACCTCTGCACTGCCCCACTGATTTGCTTCG 2677
DB 3432 CCCTTGTGTATCGGAGGGGTGGGCAACAAACCTTGTCTGCCCCCACTGATTTGCTTCG 3491
QY 2678 CAAGCATCCGAGCGCCACATATCTTCGTCGCGCTCCGCTCCCTGATCACACCCAGGTG 2737
DB 3492 CAAACATCCGAGCGCCACATATCTTCGTCGCGCTCCGCTCCCTGATATACCCAGGTG 3551
QY 2738 CCTGTGCGACTACCCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 2797
DB 3552 CATGTGCGACTACCCCGTATAGCTTTGGCATTTATCTTGTACTATCAATTACACCATATT 3611
QY 2798 TAAATCAGGATGTACGTGGAGGGGTGAAACACAGCGTGGAAAGTGCCTGCAACTGGAC 2857
DB 3612 CAAAGTCAGGATGTACGTGGAGGGGTGAGACACAGCGTGGAAAGCGCTGCAACTGGAC 3671
QY 2858 GGGGCGGAACTTGCATCTGAGATAGGAGCAGTCCGAG---ATCGATATGGAGAA 2914
DB 3672 GCGGGCGGAACTTGCATCTGAGATAGGAGCAGTCCGAGCTCCGAGCTCAGCCCATTTGCTGT 3731
QY 2915 CATCATCAGGATTCCTAGGACCCCTGCTCGTGTGTACAGCGGGGTGTTTCTTTGTTGAC 2974
DB 3732 GTCCACCAACAGTGGCAGGTCTTCCGTGTTCTTTCAGACCCCTGCCAGCTTTGTTCCAC 3791
QY 2975 AAGAACTCTCAATAACCGCAGAGTCTAG----- 3003
DB 3792 CGGCTCATCCACCTCCACAGAACATTTGTGAGACGTGACAGTACTTGTACGGGGTGGGTC 3851
QY 3004 -----ACT 3006
DB 3852 AAGCATCGCTCCTGGGCCAATTAAAGTGGAGTACGTGTTCTCTCTGTTCTTCTGTTGC 3911
QY 3007 CGTGTGGAGCTTCTCTCAATTTTCTAGGGGATCTCCCGTGTGTCTTGGCAAAATTCGC 3066
DB 3912 AGACGCGCGCTGCTCTCTGTTGTGGATGATGTTACTCATATCCAGCGGAGGCGGC 3971
QY 3067 AGTCCCAACCTCCAACTACTCA-----C 3090
DB 3972 TTTGGAGAACCTCGTAATTAATCAATGACAGATCCCTGGCGGGAACGACGCTCTGTGTC 4031
QY 3091 CAACCTCTCTCTCAATTTGTCCTGTTATCGCTGG--ATGTGTCCTGGGGGTTTAT 3148
DB 4032 CTTCTCTGTTTCTTCTGCTTTGCGTGTATCTGAAGGGGTAGGTGGTGGCGCGGCGGT 4091

Db	6252	TT	CAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGTCAAAAA	6311	Db	1	TC	CGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGACGTCCCGAGACGGTCA	60
Qy	5240	GC	-GGTTAGCTCCTCGGTCTCCGATCCGATCGTGTGTCAGAACTGTTGGCCGACGTGTTATC	5298	Qy	61	CAG	CTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCGTAGGGCGCGCTCAGCGGGTG	120
Db	6312	GC	GGTTAGCTCCTCGGTCTCCGATCGTGTGTCAGAACTGTTGGCCGACGTGTTATC	6371	Db	61	CAG	CTTGTCTGTAAAGCGGATGCCGGAGCAGACAAGCCGTAGGGCGCGCTCAGCGGGTG	120
Qy	5299	ACT	CATGGTTATGGCAGCACTGCATAACTCTTCTTACTGTGTCATGCCCATCGTAAAGATGCTT	5358	Qy	121	TT	GGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGC	180
Db	6372	ACT	CATGGTTATGGCAGCACTGCATAACTCTTCTTACTGTGTCATGCCCATCGTAAAGATGCTT	6431	Db	121	TT	GGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGAGTGC	180
Qy	5359	TT	CTGTGACTGGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGCACCGAG	5418	Qy	181	AC	CATATGAAGCTTTTGTGCAAAAGCCTAGGCCTCCAAAAAGCCTCTCACTACTTCTGG	240
Db	6432	TT	CTGTGACTGGTACTCAACCAAGTCAATCTGAGATAGTGTATGCGGCACCGAG	6491	Db	181	AC	CATATGAAGCTTTTGTGCAAAAGCCTAGGCCTCCAAAAAGCCTCTCACTACTTCTGG	240
Qy	5419	TT	CTCTTGGCCGGCGTCAATACGGGATATACCGGCCACATAGCAGAACTTTTAAAGT	5478	Qy	241	AA	TAGCTCAGAGCCGAGCGGCTCGGCTCTGTCATAATAAAAAAATAGTCAGCCA	300
Db	6492	TT	CTCTTGGCCGGCGTCAATACGGGATATACCGGCCACATAGCAGAACTTTTAAAGT	6551	Db	208	GA	TGGTAGGAGAAATACCGCATCAG-----TGTGAATACCGCACA	207
Qy	5479	GCT	CATCATTTGGAAGAGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAG	5538	Qy	301	TT	GGCGGAGAAATGGCGGAAACTGGCGGGGAGGGAATATTGGCTATTGGCCATTGTCAT	360
Db	6552	GCT	CATCATTTGGAAGAGTTCTTTCGGGGCGAAACTCTCAAGGATCTTACCGCTGTTGAG	6611	Db	236	-----	ATTGGCTATTGGCCATTGTCAT	256
Qy	5539	AT	CCAGTTGATGTAACCACTCGTGCAACCACTGATCTTACGATCTTTTACTTTTAC	5598	Qy	361	AG	TTGTATCTATATCATATATGATACATTTATTTGGCTCATCTCCATATGACGCCCA	420
Db	6612	AT	CCAGTTGATGTAACCACTCGTGCAACCACTGATCTTACGATCTTTTACTTTTAC	6671	Db	257	AG	TTGTATCTATATCATATATGATACATTTATTTGGCTCATCTCCATATGACGCCCA	316
Qy	5599	CAG	CTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGCG	5658	Qy	421	TG	TTGACATTGATTATTGACTAGTATTATAGTAATCAATTACGGGGTCAATTAGTTTCA	480
Db	6672	CAG	CTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGCG	6731	Db	317	TG	TTGACATTGATTATTGACTAGTATTATAGTAATCAATTACGGGGTCAATTAGTTTCA	376
Qy	5659	GAC	CGAAATGTTGAATCTCATACTCTTCTTTTTCATATTTTGAAGCACTTTATCA	5718	Qy	481	AG	CCCATATATGGAGTTCGCGCTTACATAAATTTACGGTAAATGGCGCGCTCGGCTGACCG	540
Db	6732	GAC	CGAAATGTTGAATCTCATACTCTTCTTTTTCATATTTTGAAGCACTTTATCA	6791	Db	377	AG	CCCATATATGGAGTTCGCGCTTACATAAATTTACGGTAAATGGCGCGCTCGGCTGACCG	436
Qy	5719	GG	TTATGTTCTATGAGCGGATACATATTTGATGTTATTTAGAAAAATTAACAATAG	5778	Qy	541	CC	CAAGCCCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGGCCAATA	600
Db	6792	GG	TTATGTTCTATGAGCGGATACATATTTGATGTTATTTAGAAAAATTAACAATAG	6851	Db	437	CC	CAAGCCCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGGCCAATA	496
Qy	5779	GG	TTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTCTAAG	5822	Qy	601	GG	ACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA	660
Db	6852	GG	TTCCGCGCACATTTCCCGAAAAAGTGCCACCTAATTTGTAAG	6895	Db	497	GG	ACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA	556
RESULT 11									
AX427812									
LOCUS									
Sequence 50 from Patent WO0232943.									
AX427812									
VERSION									
AX427812.1 GI:21537899									
KEYWORDS									
synthetic construct									
synthetic construct									
other sequences; artificial sequences.									
1									
REFERENCE									
Huang, Y. and Nabel, G.J.									
Modifications of hiv env, gag, and pol enhance immunogenicity for									
genetic immunization									
Patent: WO 0232943-A 50 25-APR-2002;									
GOVERNMENT OF THE UNITED STATES (US)									
Location/Qualifiers									
1. 6577									
/organism="synthetic construct"									
/mol type="unassigned DNA"									
/db_xref="taxon:32630"									
/note="plasmid pVR1012x/s containing HIV genes"									
ORIGIN									
Query Match 38.9%; Score 2290.6; DB 6; Length 6577;									
Best Local Similarity 68.9%; Pred. No. 0;									
Matches 3636; Conservative 0; Mismatches 1139; Indels 502; Gaps 16;									
1 TC									
Qy	1	TC	CGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGACGTCCCGAGACGGTCA	60	Qy	977	CG	ATCAGCCCTCCCGCGCGGGAAACGGTGTCAATTGGAAACGCGGATTTCCCGGTGCCAAGAT	1036

1141 GACGTAAGTACCGCCTATAGACTCTATAGGCAACCCCTTTGGCTCTTATGCAATGCTATA 1200
1037 GACGTAAGTACCGCCTATAGACTCTATAGGCAACCCCTTTGGCTCTTATGCAATGCTATA 1096
1201 CTGTTTTTGGCTTGGGCTTATACACCCCGC--TCCTTATGCTATAGGTGATGATAGC 1259
1097 CTGTTTTTGGCTTGGGCTTATACACCCCGCTTCTTATGCTATAGGTGATGATAGC 1156
1260 TTAGCCTATAGGTGTGGTTATTGACCATTTATTGACCACTCCCTCTATTGGTGACGATACT 1319
1157 TTAGCCTATAGGTGTGGTTATTGACCATTTATTGACCACTCCCTCTATTGGTGACGATACT 1216
1320 TTCCATTAATCAATTCATTAACATGGCTCTTTGGCCACAACATATCTCTATTGGCTATATGCCA 1379
1217 TTCCATTAATCAATTCATTAACATGGCTCTTTGGCCACAACATATCTCTATTGGCTATATGCCA 1276
1380 ATACTCTGCTTTCAGAGACTGACCGGACTCTGTATTTTACAGATGGGGT-CCATTT 1438
1277 ATACTCTGCTTTCAGAGACTGACCGGACTCTGTATTTTACAGATGGGGTCCCATT 1336
1439 ATTATTTACAAATTCATATACAAACGCGCTCCCGGTGCGCCGAGTTTTTTATTA 1498
1337 ATTATTTACAAATTCATATACAAACGCGCTCCCGGTGCGCCGAGTTTTTTATTA 1396
1499 CATAGCTGGGATCTCC---GACATCTCGGTTAGCTGTTCCGACATGGGCTCTCTCCG 1555
1397 CATAGCTGGGATCTCCACGCAATCTCGGGTACGTGTTCCGACATGGGCTCTCTCCG 1456
1556 GTAGCGCGGAGCTCCACATCCGAGCCCTGCTCCATCCGTCAGCGCTCATGTGCG 1615
1457 GTAGCGCGGAGCTCCACATCCGAGCCCTGCTCCATCCGCTCCAGCGCTCATGTGCG 1516
1616 TCGGCAGCTCTTGTCTTAACAGTGGAGGCCAGACTTAGGCACAGCAAAATGCCACCA 1675
1517 TCGCAGCTCTTGTCTTAACAGTGGAGGCCAGACTTAGGCACAGCAAAATGCCACCA 1576
1676 CCAACAGTGTGCCGCAAGAGCCGTCGGTAGGGTATGTGTCTGAAATAGACTCGAG 1735
1577 CCAACAGTGTGCCGCAAGAGCCGTCGGTAGGGTATGTGTCTGAAATAGACTCGAG 1636
1736 ATTGGGCTCGCACTG-GACGCAGATGGAAGACTTAAGCAGCGGCAGAGAGATGCAG 1794
1637 ATTGGGCTCGCACTGTCGCAAGTGAAGACTTAAGCAGCGGCAGAGAGATGCAG 1696
1795 GCAGCTGAGTGTGTATTCTGATAAGAGTCAGAGGTAACTCCGTTGGCGTCTGTAA 1854
1697 GCAGCTGAGTGTGTATTCTGATAAGAGTCAGAGGTAACTCCGTTGGCGTCTGTAA 1756
1855 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGTCGCGCGCGGCACAGACATA 1914
1757 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGTCGCGCGCGGCACAGACATA 1816
1915 ATAGCTGACAGACTAAACAGACTGTTCTTTCCATGGTCTTTTCTGCAGTCACCGTCGTC 1974
1817 ATAGCTGACAGACTAAACAGACTGTTCTTTCCATGGTCTTTTCTGCAGTCACCGTCGTC 1876
1975 GACGAATTCGAAGCA--ATCATGGATGCAATGAAGAGGGCTCTGCTGTGCTGCTGC 2031
1877 GACACGTGTGATCAGATATCGCGCGCTCTAGACACCATGCGCGGTGAAGGAGATACC 1936
2032 TGTGTGAGGAGCTGTTGCTTTGCCCGACGCTAGCGAACCACCGTACCGGGGGAGTG 2091
1937 AGCACCTGTGGCGCTGGGGCTGGCTGGGGCCACCATGCTGCTGGGCATGCTGATGATCT 1996
2092 CCGGCCACACTGTGCTGGAATTTGTTAGCTCTCGCACCGCGCCACAGCAACGTC 2151
1997 GCAGCGCCACCGAGGC-----TAGCAGCTTGAAGTGCACCGACTTGAAGAACGACA 2047
2152 AGCTGATCAACACCAACGCGGAGTTGGCACTCAATAGCAGCGCCCTGAACTGCAATGATA 2211
2048 CCAACACCAACGAGCAGCGCGCGCATGATCATGAGAGAGGCGGAGATCAAGAACTGCA 2107

2212 GCCTCAACACCGCTGTTGGCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCT 2271
2108 GCTTCAACATACGACACCAAGCATCCGCGCAAGGTGCAAGAGGATACGGCTTCTTCTACA 2167
2272 GTCTCTGAGAGGCTAGCAGCTGCCAGCCCTTTACCGATTTTACCAGGGCTTGGGGCCCTA 2331
2168 AGTGGACATCATCCCATCGACACGACACCAACAGCTACAGCCTGACCACTGCAACA 2227
2332 TCAGTTATGCCAAGGAGCGGCCGACGAGCCCTTACTGTGGCACTACCCCCCA- 2390
2228 CCAGCGTGTATCCACCGAGGCTGCTCCAGGTGAGCTTCGAGCCCATCCCAACCACTACT 2287
2391 --AAACCTTGGCGTATTGTCGCGGAGAGTGTGTGGTCCGCTATATGCTTCACTC 2448
2288 GGGCCCCCGCGCTTCGCGCATCTCTGAAGTCAAGGACAAGAGTTCAACGGCAAGGGCC 2347
2449 CCAGCCCCCGTGTG---GTGGGAACGACCGACAGTCCGGGGCGGCCCACTACAGCTGG 2504
2348 CTGCAACCAACGTGAGCACCCTGCACTGCAACCAACGCACTCCGCCCGGTGAGCACCC 2407
2505 GGTGAAATGATACGAGCGTCTTCTGCTTTAAACAATACAGGCCACCGCTGGGCAATTGG 2564
2408 AGCTGCTGCTGAACGCTAGCTGGCGAGGAGGAGTGTGATCCGACGCGCTAACTTCG 2467
2565 TTGCGTTGTACTGATGAATCAACTGATTCACCAAGTTCACCAAGTGTGCGAGCGCT--CCITT 2622
2468 CCGACAACCCCAAGGTGATCATCTGTCAGCTGAACGAGAGCGTGGAGATCAACTGCACCC 2527
2623 GTGTCACTCGGAGGGCGGGCAACCAACCCCTGCACTGCCCCCACTGATTTGTTCCGCAAGC 2682
2528 GCCCAACAACAACACCCGCAAGACATCCACATCGGCCCGCGCGCTTCTACACCA 2587
2683 ATCCGAGCCCAACATACTCTCGGTGGGCTCCGGTCCCTGATACACCCAGTGTGCTGG 2742
2588 CCGCGAGATCATCGCGACATCCGCGAGGCCCACTGCAACCTGAGCCGCGCAAGTGA 2647
2743 TGCACTACCGGTATAGGCTTTGCGATTATCTTGTACCATCAACTACACCATATTTAAAA 2802
2648 AGCACACCTGAACAAGATCGTGATCAAGCTGCGGAGCAGTTTCGGCAACAAGACCATCG 2707
2803 TCAGGATGTACTGAGGGGTGGAACACAGAGCTGGAAGCTGCTGCAACTGACCGCGG 2862
2708 TGTTCAGACACAGCAGCGCGCGGACCCCGAGATCGTGACCCACACACTCACTCGCGG 2767
2863 GCGAAGTTGGATCTGGAAGATAGGGAAGGTCCGAGATCGATATGGAAGACATCACAT 2922
2768 GCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCACAGCAGCCTGGTTCAACAGCACCT 2827
2923 CAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTCTTTGTGACAAGAATCC 2982
2828 GGAGCACCGAGGCGAGCAACAAC-----ACGAGGGCAGCGACCATCACTCCTGCCCT 2981
2983 TCACAATACCGCAGAGTCTAGACTCGTGGTGAATCTCTCAATTTTCTAGGGGATCTC 3042
2882 GCGCATCAAGCAGATCATCAACATGTGCGAAGGTGGCAAGGCCATGTACGCCCCC 2941
3043 CGGTGTCTTGGCCAAAATTGCGAGTCCCAACCTTCAATCACTCAACCACTCTCTGTC 3102
2942 CCATAGCGCGCAGATCCGCTGCGAGCAACAATCAACCGGCTGCTGCTGACCGCGACG 3001
3103 CTCCAATTTGCTCTGTTATCGCTGATGTGTCGGCGGTTTTTATCATATATCTCTTCA 3162
3002 GGGCAACAAGCAACAACGAGAGGAGATCTTTCGTCGGCGCGCGGAGCATCGCGGACA 3061
3163 TCTGCTGTATGCTCTCATCTTCTTATTTGTTCTTCTGGATTATC-----AAGTATGTTG 3218
3062 ACTGGCGCAGCGAGCTGTACAAGTCAAGTGTGTAAGATCGAGCCCTTGGCGTGGCCC 3121
3219 CCGGTTTGTCTCTAATTCAGGATCAACAACAGCAGTACGAGACCATGCAAAACCTGTC 3278
3122 CCACCAAGGCCAAGCTTACCGTCCAGGCCCGCAGCTGCTGAGCGGCATCGTGCAGCAGC 3181
3279 ACGACTCTGCTCAAGGCCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACTACGGAT 3338

JOURNAL	Patent: WO 0232943-A 15 25-APR-2002;
FEATURES	GOVERNMENT OF THE UNITED STATES (US)
source	Location/Qualifiers
	1. .6505
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="plasmid pVR1012x/s containing HIV genes"
ORIGIN	
	Query Match 38.8%; Score 2283.4; DB 6; Length 6505;
	Best Local Similarity 69.5%; Pred. No. 0;
	Matches 3656; Conservative 0; Mismatches 1061; Indels 546; Gaps 19;
QY	1 TCSCGGCTTCGCTGATGACGGTGAAACCTCTGACACATGCAAGTCCCGAGACGGTCA 60
DB	1 TCSCGGCTTCGCTGATGACGGTGAAACCTCTGACACATGCAAGTCCCGAGACGGTCA 60
QY	61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAAGGGCGGTGAGCGGTG 120
DB	61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAAGGGCGGTGAGCGGTG 120
QY	121 TTGGCGGGTCTCGGGGCTGCTTAACTATCGCGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB	121 TTGGCGGGTCTCGGGGCTGCTTAACTATCGCGCATCAGAGCAGATTGTACTGAGAGTGC 180
QY	181 ACCATATGAAGCTTTTTCGAAAGCCCTGAGCCCTCCAAAAGCCCTCCTCACTACTTCTGG 240
DB	181 ACCATATGCGG-----TGGAATAATCCGCACA 207
QY	241 AATAGCTCAGAGCGCGGCGCTCGGCTCTGCTGATAATAAAAAAATAGTCAGCCA 300
DB	208 GATGCGTAAGGAGAAAATACCGCATCAG----- 235
QY	301 TGGGGCGGAGAAATGGGCGGAACTGGGCGGGAGGGAATATTGGCTATTGGCCATTGCAT 360
DB	236 -----ATTGGCTATTGGCCATTGCAT 256
QY	361 ACSTTGTATCTATACATAATATGTACATTTATATTGGCTCATGTCCAATATACCGCCA 420
DB	257 ACSTTGTATCTATACATAATATGTACATTTATATTGGCTCATGTCCAATATACCGCCA 316
QY	421 TGTGACATTTGATTTGACTAGTATTATAGTAATCAATTAACGGGTCTAGTTAGTTTAT 480
DB	317 TGTGACATTTGATTTGACTAGTATTATAGTAATCAATTAACGGGTCTAGTTAGTTTAT 376
QY	481 AGCCCATATATGAGTTTCCCGCTTACATACTTACGGTAAATGGCCCGCTGGCTGACCG 540
DB	377 AGCCCATATATGAGTTTCCCGCTTACATACTTACGGTAAATGGCCCGCTGGCTGACCG 436
QY	541 CCACAGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAACGCAATA 600
DB	437 CCACAGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAACGCAATA 496
QY	601 GGGACTTCCATTTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCATTTGGCAGTA 660
DB	497 GGGACTTCCATTTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCATTTGGCAGTA 556
QY	661 CATCAAGTGTATCATATGCGCAAGTCCCGCCCTTATTGACGTCAATGACGGTAAATGGCCC 720
DB	557 CATCAAGTGTATCATATGCGCAAGTACGCCCTTATTGACGTCAATGACGGTAAATGGCCC 616
QY	721 GCCTGGCATTTGCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAC 780
DB	617 GCCTGGCATTTGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC 676
QY	781 GTATTAGTCACTGCTATACCATGGGTGATGGGTGAGTATGCGGTAACTGCCCATTTGGCAGTA 840
DB	677 GTATTAGTCACTGCTATACCATGGGTGATGGGTGAGTATGCGGTAACTGCCCATTTGGCAGTA 736
QY	841 TAGCGGTTTGACTCACGGGGAATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900
DB	737 TAGCGGTTTGACTCACGGGGAATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 796
QY	901 TTTTGGCACAATAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGCTTGACG 960
DB	797 TTTTGGCACAATAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGCTTGACG 856
QY	961 CAAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAGCAGAGCTCGTTTGTAGTAAC 1020
DB	857 CAAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAGCAGAGCTCGTTTGTAGTAAC 916
QY	1021 CGTCAGATCGCTTGGAGAGCCCATCCACGCTGTTTTGGACTCCCATAGAAAGACACCGGGAC 1080
DB	917 CGTCAGATCGCTTGGAGAGCCCATCCACGCTGTTTTGGACTCCCATAGAAAGACACCGGGAC 976
QY	1081 CGATCAGAGCTTCGCGCGCGGGAACGGTGCATTTGGAACCGGATTTCCCGTGCACAGAT 1140
DB	977 CGATCAGAGCTTCGCGCGCGGGAACGGTGCATTTGGAACCGGATTTCCCGTGCACAGAT 1036
QY	1141 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1200
DB	1037 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1096
QY	1201 CTGTTTTTGGCTTTGGGCGCTATACACCCCGGC--TCCTTATGCTATAGGTGATGTTATAGC 1259
DB	1097 CTGTTTTTGGCTTTGGGCGCTATACACCCCGCTTCTTATGCTATAGGTGATGTTATAGC 1156
QY	1260 TTAGCCTATAGGTGTGGGTATTGACCAATATTGACCACTCCCTTATTTGGTGAAGTACT 1319
DB	1157 TTAGCCTATAGGTGTGGGTATTGACCAATATTGACCACTCCCTTATTTGGTGAAGTACT 1216
QY	1320 TTCCATTACTAATCCATAACATGGCTTTTGGCACAACATCTCTATTGGCTATATGCCA 1379
DB	1217 TTCCATTACTAATCCATAACATGGCTTTTGGCACAACATCTCTATTGGCTATATGCCA 1276
QY	1380 ATACTCTGCTCTTACAGAGCTGACCGACTCTGTATTTTACAGATGGGT--CCATTT 1438
DB	1277 ATACTCTGCTCTTACAGAGCTGACCGACTCTGTATTTTACAGATGGGTGCCATTT 1336
QY	1439 ATTATTTACAAATTTACATATACAAACCGCGTCCCGTGGCCCGCAGTTTTTATATAA 1498
DB	1337 ATTATTTACAAATTTACATATACAAACCGCGTCCCGTGGCCCGCAGTTTTTATATAA 1396
QY	1499 CATAGCTGGGATCTCC--GACATCTCGGTACGTTTCCGGACATGGCTCTTCTCCG 1555
DB	1397 CATAGCTGGGATCTCCAGCGAATCTCGGTACGHTTCCGGACATGGCTCTTCTCCG 1456
QY	1556 GTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCACGCGCTCATGTGCG 1615
DB	1457 GTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCACGCGCTCATGTGCG 1516
QY	1616 TCGGAGCTCTTGTCTCTTAAAGTGGAGCCAGACTTAGGCACAGACAAATGCCACCA 1675
DB	1517 TCGGAGCTCTTGTCTCTTAAAGTGGAGCCAGACTTAGGCACAGACAAATGCCACCA 1576
QY	1676 CCACAGTGTGCCCAAGGCGTGGCGTAGGGTATGTCTGTAATAATGAGCTCGGAG 1735
DB	1577 CCACAGTGTGCCCAAGGCGTGGCGTAGGGTATGTCTGTAATAATGAGCTCGGAG 1636
QY	1736 ATTGGGCTCGCAC--CTGGACGCGAGATGGAAGACTTTAAGCAGCGCGCAGAAAGATGCGAG 1794
DB	1637 ATTGGGCTCGCACGCGGCTGACGAGATGGAAGACTTTAAGCAGCGCGCAGAAAGATGCGAG 1696
QY	1795 GCGCTGAGTGTGTGTTCTGATTAAGAGTCAGAGTAACTCCGTTGGTGGTCTGTAA 1854
DB	1697 GCGCTGAGTGTGTGTTCTGATTAAGAGTCAGAGTAACTCCGTTGGTGGTCTGTAA 1756
QY	1855 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTCGCGCGCGCCACCCAGACATA 1914
DB	1757 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTCGCGCGCGCCACCCAGACATA 1816
QY	1915 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGTCTTTTCTGAGTCAACCGTCGTC 1974
DB	1817 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGTCTTTTCTGAGTCAACCGTCGTC 1876

Db 3964 AGAAATTAAGCAAGTAGCTATTAAAGTCAGAGGGAGAGAAAATGCGCTCCAAATGTG 4023

Qy 3879 -----GCTCTTCCGCTTC 3891

Db 4024 AGGAAGTAATAGAGAAATCATAGAAATTTTAAAGCCATCATGCCCTTAATCTTCCGCTTC 4083

Qy 3892 CTGCTCAGTACTCGTTCGGCTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC 3951

Db 4084 CTCGCTCAGTACTCGTTCGGCTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC 4143

Qy 3952 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGGAAAGAAACATGTGAGC 4011

Db 4144 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGGAAAGAAACATGTGAGC 4203

Qy 4012 AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAAGCGCGTTCCTGCTGCGGTTTTTCCATAG 4071

Db 4204 AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAAGCGCGTTCCTGCTGCGGTTTTTCCATAG 4263

Qy 4072 GCTCCGCCCCCTGACGAGCATCACAAAATCAGCGTCAAGTCAGAGTGGCGAAACCC 4131

Db 4264 GCTCCGCCCCCTGACGAGCATCACAAAATCAGCGTCAAGTCAGAGTGGCGAAACCC 4323

Qy 4132 GACAGACTATAAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGT 4191

Db 4324 GACAGACTATAAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGTCTCCTGT 4383

Qy 4192 TCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTGGCGCT 4251

Db 4384 TCCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTGGCGCT 4443

Qy 4252 TTCTCAATGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGGG 4311

Db 4444 TTCTCAATGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGCTGGG 4503

Qy 4312 CTGTGTGACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCGGTAACTATCGTCT 4371

Db 4504 CTGTGTGACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCGGTAACTATCGTCT 4563

Qy 4372 TGAGTCCAAACCCGGTAAGACACAGCTTATGCCACTGGCAGCAGCCAATGGTAAACAGGAT 4431

Db 4564 TGAGTCCAAACCCGGTAAGACACAGCTTATGCCACTGGCAGCAGCCAATGGTAAACAGGAT 4623

Qy 4432 TACGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACCTACCG 4491

Db 4624 TACGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACCTACCG 4683

Qy 4492 CTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGGAAA 4551

Db 4684 CTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCTTCGGAAA 4743

Qy 4552 AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTGT 4611

Db 4744 AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTGT 4803

Qy 4612 TTGCAAGCAGAGTATACGGCGAGAAAAGAGTCTCAGAGAGATCCTTTGATCTTTTC 4671

Db 4804 TTGCAAGCAGAGTATACGGCGAGAAAAGAGTCTCAGAGAGATCCTTTGATCTTTTC 4863

Qy 4672 TACGGGTCTGAGCTCAGTGGAAACGAAAACCTCAGCTTAAAGGATTTTGGTCAATGAGATT 4731

Db 4864 TACGGGTCTGAGCTCAGTGGAAACGAAAACCTCAGCTTAAAGGATTTTGGTCAATGAGATT 4923

Qy 4732 ATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGATTTTAAATCAATCTA 4791

Db 4924 ATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTAATAAGATTTTAAATCAATCTA 4983

Qy 4792 AAGTATATAGTAACTTTGGTCTGACAGTTACCAATGCTTAAATCAGTGGAGCACCTAT 4851

Db 4984 AAGTATATAGTAACTTTGGTCTGACAGTTACCAATGCTTAAATCAGTGGAGCACCTAT 5043

Qy 4852 CTGAGGATCTGTCTATTTTCGTTTCATCATAGTTGCTGACTC 4894

Db 5044 CTGAGGATCTGTCTATTTTCGTTTCATCATAGTTGCTGACTC 5086

RESULT 13

AX427796

LOCUS AX427796 6460 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 34 from Patent WO0232943.

ACCESSION AX427796

VERSION AX427796.1 GI:21537883

KEYWORDS synthetic construct

SOURCE other sequences; artificial sequences.

ORGANISM 1

REFERENCE 1

AUTHORS Huang, Y. and Nabel, G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization

JOURNAL Patent: WO 0232943-A 34 25-APR-2002; GOVERNMENT OF THE UNITED STATES (US)

FEATURES

Location/Qualifiers

1..6460

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="plasmid pVR1012x/s containing HIV genes"

ORIGIN

Query Match 38.8%; Score 2283; DB 6; Length 6460;

Best Local Similarity 69.9%; Pred. No. 0;

Mismatches 3595; Conservative 0; Mismatches 1195; Indels 355; Gaps 24;

Qy 1 TCGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGCACTCCCGGAGACGGTCA 60

Db 1 TCGCGGTTTCGGTGTATGACGGTGAACCTCTGACACATGCACTCCCGGAGACGGTCA 60

Qy 61 CAGCTTGTCTGTAAAGCGGATGCGGGAGCAGACAAGCCGTCAGGCGCGTCAAGCGGGTG 120

Db 61 CAGCTTGTCTGTAAAGCGGATGCGGGAGCAGACAAGCCGTCAGGCGCGTCAAGCGGGTG 120

Qy 121 TTGGCGGGTGTTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC 180

Db 121 TTGGCGGGTGTTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC 180

Qy 181 ACCATATGAAGCTTTTTCGAAAAGCCTAGGCCCTCCAAAAAGCCTCCTCACTACTTCTGG 240

Db 181 ACCATATGAAGCTTTTTCGAAAAGCCTAGGCCCTCCAAAAAGCCTCCTCACTACTTCTGG 240

Qy 241 AATAGCTCAGAGCGCGGCGGCTCGGCTCTGCTATAATAAAAAAATTTAGTCAGCCA 300

Db 208 GATGCGTAAGGAGAAAATACCGCATCAG----- 235

Qy 301 TGGGCGGAGAAATGGGCGGAACTGGGCGGGAGGGAATTTATTGGCTATTGGCCATTGGCAT 360

Db 236 -----ATTGGCTATTGGCCATTGGCAT 256

Qy 361 ACGTTGTATCTATCATATATGTACATTTATTATTGGCTCATGTCCAATATGACGCCCA 420

Db 257 ACGTTGTATCTATCATATATGTACATTTATTATTGGCTCATGTCCAATATGACGCCCA 316

Qy 421 TGTTCAGATTTGATTTTGTACTAGTTTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT 480

Db 317 TGTTCAGATTTGATTTTGTACTAGTTTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT 376

Qy 481 AGCCCATATATGAGGTTTCGGGTTACATACTTACCGGTAATGGCCCGCTGGCTGACCG 540

Db 377 AGCCCATATATGAGGTTTCGGGTTACATACTTACCGGTAATGGCCCGCTGGCTGACCG 436

Qy 541 CCCAAGCAGCCCGCCCATTTGAGCTCAATAATGACCTATGTTCCCATAGTAAGCCCAATA 600

Db 437 CCCAAGCAGCCCGCCCATTTGAGCTCAATAATGACCTATGTTCCCATAGTAAGCCCAATA 496

Qy 601 GGGACTTTTCCATTTGACGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 660

Db 497 GGGACTTTTCCATTTGACGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 556

Db	2717	CAAGACCATCGTGTTCAGACACAGCAGCGCGCGGACCCCGAGATCGTGACCCACAGCTT	2776
Qy	2738	CTGTGTCGATACCCGTATAGGCTTTGGCATATTCCTTGTACCATCAACTACACCATATT	2797
Db	2777	CAACTGGCGGCGAGTCTTCTTACTGCAACAGCGCTAGCCTGCCCTGCGCATCAAGCA	2836
Qy	2798	TAAATACAGGATGTAAGTGGAGGGGTGGAACACAGCGTGGAAGTGCCTTGCAACTGGAC	2857
Db	2837	GATCATCAATGTGGCAGAAAGTGGGCAAGGCATGTACGCCGCCGCCCATCAGCGGCCA	2896
Qy	2858	GCGGGGGAAGTTCGATCTGGAA-----GATAGGAGACAGTCCGA	2899
Db	2897	GATCCGCTGCAGCAGCAACAATCACCGGCCCTGTGCTGACCCGCGACGGGGCAACAGCAA	2956
Qy	2900	GATCGATATGGAGAACATCACATCAGGATTCCTTAGGACCCCTCTCGTGTATACGGCGGG	2959
Db	2957	CAACGAGCGAGATCTTCGTCGGGCGGCGGACATCGCGGCAAACTGGCGCAGCGA	3016
Qy	2960	GTTTTTCTTGTGACAAGAAATCTCAATAACGACAGAGTCTAGACTGTGTGTGGACTTC	3019
Db	3017	GCTGTACAAGTACAGGTGTGAAGATCGAGCCCTCGGGCGTGGCCGCCACCAAGGCCAA	3076
Qy	3020	TCTCAATTTCTAG-----GGGATCTCCGCTGTCTTGGCCAAAATTCGAGT	3069
Db	3077	GCTTACCGTCCAGGCCCGCAGCTGTGTAGCGGCATCGTGACAGCAGCAAGAACCTGCT	3136
Qy	3070	CCCCAACCTCAATCACTACCAACCTCTGTCTCTCAATTTGTCTCTGTTATCGCTGGA	3129
Db	3137	GCGGCCATCGAGGCCCAGCAGACCTGCTGCAGCTGACCGTGTGGGCATCAAGCAGCT	3196
Qy	3130	-----TGTGCTGCGCGGTTTATCATATTTCTCTTTCATC-----	3164
Db	3197	GCAGGCCCGCACCTCGCGCTGAGCGGTACCTGGAAGGACGACGAGCAGCTCGAGCAGAT	3256
Qy	3165	CTGCTGCTATGCTCATCTCTTATTATGTTCTTCTGATTTATCAAGTATGTTGCCGTT	3224
Db	3257	CTGGAACCAACACCATCTGGATGGAGTGGGACCGCGAGATCAACAACTACACAGCCCTGAT	3316
Qy	3225	TGTCTCTTAATTCAG--GATCAACAACACAGTACGGGACCATGCAAAACCTGCACGA	3282
Db	3317	CCACAGCTGATCGAGGAGGCGAGAACCCAGCAGGAGAGAAAGCAGCAGGAGCTGTGGA	3376
Qy	3283	CTCCTGCTCAAG-----TGCTGTGCGCGGTTTATCATATTTCTCTTTCATC-----	3319
Db	3377	GCTGGAACAAGTGGGCCAGCCTGTGGAATCTGGTTCAACATCACCAACTGGCTGTGTGAGG	3436
Qy	3320	CTGTACAAAACCTACGATGGAATTTGC---ACCTGTATTCCCATCCCATCGTCTGGGC	3376
Db	3437	ATCCAGATCTGCTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCGCCGTGCC	3496
Qy	3377	TTTCGCAAAATACCTATGGAGTGGGCCTCAGTCCGTTTCT-CTTGGCTCAGTTTACTAG	3435
Db	3497	TTCTTTGACCTGGAAGTGCCACTCCCACTGTCTTCTTAATAAAATGAGGAAATTC	3556
Qy	3436	TGCCATTTGTTAGTGGTTCGTAAGGGTTTCCCCCACTGTTTGGCTTTTCAAGTATATGGA	3495
Db	3557	ATCGCATTTGTCTAGTAGTGTCTATTCTATTCTGGGGGGTGGGGTGGGCGGACAGCAA	3616
Qy	3496	TGATGTGTTATTTGGGGCCAAAGTCTGTACAGCATCGTGAGTCCCTTTATACCGCTGTAC	3555
Db	3617	GGGGGAGGATTTGGGAAGACAATAAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGGTAC	3676
Qy	3556	CAATTTTCT-----TTTGTCTCTGGGTATACATTTAAGAATTCAGACTC---	3599
Db	3677	CCAGGTGCTGAAGAAATTGACCCGGTTCTCTCGGCCAGAAAGCAGGCACATCCCT	3736
Qy	3600	---GAGCAAGTCTAGAAAGCGCGCCCAAGATATCAAGGATCCACTAGCGGTTAGAGCTCG	3656
Db	3737	TCTCTGTGACACACCTGTCTCACGCGCCCTGGTTCTTAGTTCCAGCCCACTCATAGGACA	3796
Qy	3657	CTGATCAGCCTCGACTGTGCTTCTAGTTGCCAGGCATCTGTTGTTGGCCCTCCCCCGT	3716

Db	3797	CTCATAGCTCAGAGGGCTCGCGCTTCAATCCACCCGCTAAAGTACTTTGGAGCGGCTC	3856
Qy	3717	GCCTTCTCTTGA-----CCTGGAAAGTGCCACTCCCACTGTCTCTTTCTTAATAATGA	3770
Db	3857	TCCCTCCCTCATCAGCCCAACCAACCTAGCCTCCAAAGAGTGGGAAGAAATTAAG	3916
Qy	3771	GGAAATTGCATCGCATTTGTCTAGTAGGTGTCTATCTATTCTCTGGGGGTGGGGTGGGCA	3830
Db	3917	CAAGTAGCTATTAAAGTGCAGAGGAGAGAAATGCTCCAAATGTGAGGAAGTAATG	3976
Qy	3831	GGACACAAAGGGGAGGATTTGGGAAGACAAATAGCAGGCATGCTGG--GGAGCTCTCCGCT	3889
Db	3977	AGGAATCATAGATTTTAAAGCCCATGATTTAAAGCCATCATGGCTTAATCTTCGCT	4036
Qy	3890	TCTCGCTCACTGACTCGCTCGCTCGTTCGTTGCGCTCGGCGAGCGGTATCAGCTCAC	3949
Db	4037	TCTCGCTCACTGACTCGCTCGCTCGTTCGTTGCGCTCGGCGAGCGGTATCAGCTCAC	4096
Qy	3950	TCAAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCAGGAAGACATGTGA	4009
Db	4097	TCAAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGCAGGAAGACATGTGA	4156
Qy	4010	GCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGGCCGCGTGTGCTGGCGTTTTCCAT	4069
Db	4157	GCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGGCCGCGTGTGCTGGCGTTTTCCAT	4216
Qy	4070	AGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC	4129
Db	4217	AGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC	4276
Qy	4130	CGACAGGACTATAAGATACGAGGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCT	4189
Db	4277	CGACAGGACTATAAGATACGAGGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCT	4336
Qy	4190	GTTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCG	4249
Db	4337	GTTTCCGACCTCGCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCG	4396
Qy	4250	CTTTCTCAATGTCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTTCGCTCCAAGCTG	4309
Db	4397	CTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTTCGCTCCAAGCTG	4456
Qy	4310	GGCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACATCTCT	4369
Db	4457	GGCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACATCTCT	4516
Qy	4370	CTTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG	4429
Db	4517	CTTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG	4576
Qy	4430	ATTAGCAGAGCAGGATGTAGCGGTGCTACAGAGTCTTGAAGTGTGGTGGCCCTAATCTAC	4489
Db	4577	ATTAGCAGAGCAGGATGTAGCGGTGCTACAGAGTCTTGAAGTGTGGTGGCCCTAATCTAC	4636
Qy	4490	GGCTACACTAGAGGACAGTATTGTTGATCTCGCTCTGCTGAAGCCAGTACTCTTCGGA	4549
Db	4637	GGCTACACTAGAGGACAGTATTGTTGATCTCGCTCTGCTGAAGCCAGTACTCTTCGGA	4596
Qy	4550	AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGGTTTTTTT	4609
Db	4697	AAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACCGCTGGTAGCGGTGGTTTTTTT	4756
Qy	4610	GTTTGAAGCAGCAGATTAACGCGAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTT	4669
Db	4757	GTTTGAAGCAGCAGATTAACGCGAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTT	4816
Qy	4670	TCTACGGGCTCTCAGCTCAGTGGACGAAACTCAGCTTTAAGGGATTTTGGTCATGAGA	4729
Db	4817	TCTACGGGCTCTCAGCTCAGTGGACGAAACTCAGCTTTAAGGGATTTTGGTCATGAGA	4876
Qy	4730	TTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTTTAATCAATC	4789
Db	4877	TTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTTTAATCAATC	4936

Db 8392 AGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGG 8451
QY 4714 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATG 4773
Db 8452 GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATG 8511
QY 4774 AGTTTTTAATCAATCTAAAGTATATATAGATGAAAATTGGTCTGACAGTTACCAATGCTT 4833
Db 8512 AAGTTTTAAATCAATCTAAAGTATATATAGTAAAATTGGTCTGACAGTTACCAATGCTT 8571
QY 4834 ATTCAGTGGGACCTATCTCAGCGATCTGCTATTTTCTGTTTCATCATAGTTGCTGACT 4893
Db 8572 AATCAGTGGGACCTATCTCAGCGATCTGCTATTTTCTGTTTCATCATAGTTGCTGACT 8631
QY 4894 CCCGTCGTGTAGATAACTACGATACGGAGGCTTACCATCTGGCCCCAGTGTCTCAAT 4953
Db 8632 CCCGTCGTGTAGATAACTACGATACGGAGGCTTACCATCTGGCCCCAGTGTCTCAAT 8691
QY 4954 GATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCAGCCGG 5013
Db 8692 GATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACACAGCAGCCGG 8751
QY 5014 AAGGGCCGAGCGCAGAGTGTCTTCACTTTTATCCGCTCCATCCAGTCTATTAATG 5073
Db 8752 AAGGGCCGAGCGCAGAGTGTCTTCACTTTTATCCGCTCCATCCAGTCTATTAATG 8811
QY 5074 TTGCCGGGAAGCTAGAGTAAGTGTGCGCAGTTAATAGTTTGGCGCAACGTTGTTGCCAT 5133
Db 8812 TTGCCGGGAAGCTAGAGTAAGTGTGCGCAGTTAATAGTTTGGCGCAACGTTGTTGCCAT 8871
QY 5134 TGCTACAGGATCGTGGTGTACGCTCGCTGCTTGTGATGGCTTCATTCAGCTCCGGTTC 5193
Db 8872 TGCTACAGGATCGTGGTGTACGCTCGCTGCTTGTGATGGCTTCATTCAGCTCCGGTTC 8931
QY 5194 CCAACCATCAAGCGAGTGTACATGATCCCGCATGTTGTGCGCAAAAGCGGTAGCTCCTT 5253
Db 8932 CCAACCATCAAGCGAGTGTACATGATCCCGCATGTTGTGCGCAAAAGCGGTAGCTCCTT 8991
QY 5254 CGTCTCCGATCGTGTGTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGC 5313
Db 8992 CGTCTCCGATCGTGTGTCAGAAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGC 9051
QY 5314 AGACATGCAATATCTTCTACTGTCATGCAATCCGTAAGATGCTTTCTGTGATCGTGA 5373
Db 9052 AGACATGCAATATCTTCTACTGTCATGCAATCCGTAAGATGCTTTCTGTGATCGTGA 9111
QY 5374 GTACTCAACCAAGTCATCTGAGAATAGTGTATGCGCGCACCGAGTTGCTTTGCCCCGGC 5433
Db 9112 GTACTCAACCAAGTCATCTGAGAATAGTGTATGCGCGCACCGAGTTGCTTTGCCCCGGC 9171
QY 5434 GTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAA 5493
Db 9172 GTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAA 9231
QY 5494 AGTTCTTTCGGGGCGAAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA 5553
Db 9232 AGTTCTTTCGGGGCGAAAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA 9291
QY 5554 ACCACTCGTGCAACCAATGATCTTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTG 5613
Db 9292 ACCACTCGTGCAACCAATGATCTTTCAGCATCTTTTACTTTTCCACGAGGTTTCTGGGTG 9351
QY 5614 AGCAAAAACAGGAAGCAAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAATGTTG 5673
Db 9352 AGCAAAAACAGGAAGCAAAATGCCCAAAAAAGGGAATAAGGGCGACACGGAATGTTG 9411
QY 5674 AATACTCATATCTTCTCTTTTCAATATATTAAGACATTTATCAGGGTTATGTTCTCAT 5733
Db 9412 AATACTCATATCTTCTCTTTTCAATATATTAAGACATTTATCAGGGTTATGTTCTCAT 9471
QY 5734 GAGCGGATACATATTTGAATGATTTTAGAAAAATAAACAATAGGGGTTCCGCGCACATT 5793

Db 9472 GAGCGGATACATATTTGAATGATTTTAGAAAAATAAACAATAGGGGTTCCGCGCACATT 9531
QY 5794 TCCCGAAAAGTGCCACCTGACGCTTAAGAAACCACTATTATTCATGACATTAACTTATAA 5853
Db 9532 TCCCGAAAAGTGCCACCTGACGCTTAAGAAACCACTATTATTCATGACATTAACTTATAA 9591
QY 5854 AATAGGCGTATCACGAGGCCCTTTCGTC 5882
Db 9592 AATAGGCGTATCACGAGGCCCTTTCGTC 9620

RESULT 15
AX164581
LOCUS AX164581
DEFINITION Sequence 3 from Patent WO0138360.
ACCESSION AX164581
VERSION AX164581.1 GI:14545515
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Coit, D.C., Medina-Selby, A.C., Selby, M.C. and Houghton, M.C.
TITLE Novel hcv non-structural polypeptide
JOURNAL Patent: WO 0138360-A 3 31-MAY-2001;
CHIRON CORPORATION (US)
FEATURES
Location/Qualifiers
1..9620
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1990..7305
/note="unnamed protein product; pDeltaNS3NS5"
/codon_start=1
/transl_table=11
/protein_id="CAC42723.1"
/db_xref="GI:14545516"
/translation="MAAFAAQGYKVLNPSVAATLGFAGYMSKAHGIDPNTRTGVRT
LTVLATATPGSVTPHPNIEVALSTGEIIFYGKAIPLVIGKRHLIFCHSKKAC
DELAALVALGINAVAYVGLDVSVPTSGDVVATDALMTGYTGDPSVDCNTCV
TOTVDSLDPFTIETILPQDASVTRQRTGRKPGIYRFVAPEPSPGMSV
LCECYDCAWYELTPAETTVRLRAYMTPGLPVQDHLFEWEGVFTGLTHDAHFLS
QTKGSENPYLVAQATVCARAAPPSPSWQDWKCLIRKLTHTGPTLRLRLGAVQ
NEITLHPVTYIMTMSADLEVTVLWGVLAALAAAYCLSTGCVVIVGRVLSG
KPAIIPDRELYREFDEMEECQHLPIEQGMMLAEQFKALGLLQTLASRQAEVIA
AVQTNWKLFTWAKHMFISGIVLAGSLTLPGNLPGNLAISLMAFTAAVTSPLTSTQ
LLFNILGWNVAALAAPGAATFVAGLAGAAGISVGLGKVLIDILAGYAGVAGALV
AFKIMSGEVPSTEDLVNLLPATLSPGALVGVCAAILRRHVPGEGAVQWNNRLIAF
ASRGNHVSPTHVPESDAARVATLSSLTVTQLRLRHQWISSECTTFCSSGWLADI
WMI CEVLSDFKTLWKLMPQLPGIPFVSCQRYGVWGRGIMHTRCHGAEITGH
VKNGTMRIVGPRTCRNMSTGTFINAYTTCPTPLPAPNTTALMRVSAEYVEIRGV
GDFHVTMTDNLKCPQVSPFETELDGLVRLHFAFPCKPLREEYSFVYGLHEY
PVGSQLPCPEBPDVALVLSMTDPGSHITAEAGRLRLARGLPSSASSASQSLAPSLK
ATCTANHDSDAELJEANLLWQEMGNITRVESENKVVILDSFDPLVAEEDREISV
PAEIRLKRFRFAQLPVWARPYNPLVETWKKPDEYPPVVGCCPLPPKSPVPPPR
KKRTVLTETSLTALAEATRSFGSSSTSGITGNTTSSPAPGCPSPDAESYS
SNMPLGEPDPLDGSWTSSEANAEDEVCCSMYSYMTGALVTPCAAAEQKLEIN
ALSNLGRHLNVLVYSTSRACORQKVTFDRLQVLDHSYQVLDKEVAAAQKVLN
LSVBACSLTPPHSAKSGFYGAKDVRCHARKAVTHNSWKDLLEDNVTPIDTITMA
KNEVCQVQPEKGRKPARLIIVFDLGVRYCEKMALYDVVTKLPLAVMGSSYGYOYSPG
QVRELVQAWKSKKTPMGFSYDTRCFDSVTESDIRTEAIIYQCCDLDPQARVALKSL
TERLVYGGTLNTRNGENGYRRCRAGVLTTCGNTLTCTIIRAKAACRAAGLQDCITML
VCGDDLVI CESAGQEDAAARAFTEAMTRYSAAPPDPPQPEYDLELITSSCSNVSV
AHDGAKRYYLTDRPTTPLARAAMETARTHPVNSWLNIIIMPATFLWARMILMTFF
SVLIARDLOEQALDCEIYGACYSIRPLDLPPIIORLHLSAFSLHSYSGEINRVAAC
LKLGVPLPRAWRHARSVRALLAGGAAICGKYLEFNVAVRTKLPLPIAAAGQLD
LSGWTAGYSGGDIYHSVSHRPRWIFWCLLLLAAGVGTYILLPNR"

Query Match 38.8%; Score 2281.8; DB 6; Length 9620;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2292; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
ORIGIN

QY 3574 GGGTATACATTTAAGAAATTCAGACTCGAGCAAGTCTAGAAAGCGCGCCCAAGATATCAAG 3633
DB 7312 GGGTAAACACTCCGGCTTAAAAAATAAATAATCTAGAAAGCGCGCCCAAGATATCAAG 7371
QY 3634 GATCCACTAGCGTTAGAGCTCGGTGATCAGCTCGACTGTGGCTTCTAGTTGCGAGCCA 3693
DB 7372 GATCCACTAGCGTTAGAGCTCGGTGATCAGCTCGACTGTGGCTTCTAGTTGCGAGCCA 7431
QY 3694 TCTGTTGTTGCCCTCCCGCTCCCTTCTTGAACCTGGAAGTGCACCTCCCACTGTC 3753
DB 7432 TCTGTTGTTGCCCTCCCGCTCCCTTCTTGAACCTGGAAGTGCACCTCCCACTGTC 7491
QY 3754 CTTTCTTAATAAATGAGGAATTCGATCGCATTTGCTGAGTAGGTGCTATTTCTTG 3813
DB 7492 CTTTCTTAATAAATGAGGAATTCGATCGCATTTGCTGAGTAGGTGCTATTTCTTG 7551
QY 3814 GGGGTGGGTGGGCGAGGACAGCAAGGGGAGGATTTGGGAAGCAATAGCAGGCAATGCT 3873
DB 7552 GGGGTGGGTGGGCGAGGACAGCAAGGGGAGGATTTGGGAAGCAATAGCAGGCAATGCT 7611
QY 3874 GGGGAGCTCTCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGGTGCTGCTGGCTGGCG 3933
DB 7612 GGGGAGCTCTCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGGTGCTGCTGGCTGGCG 7671
QY 3934 AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCACAGAATCAGGGGATAACGC 3993
DB 7672 AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCACAGAATCAGGGGATAACGC 7731
QY 3994 AGGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGAAACCGTAAAAAGCCGCGTT 4053
DB 7732 AGGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGAAACCGTAAAAAGCCGCGTT 7791
QY 4054 GCTGGCGCTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAATCGAGCTCAAG 4113
DB 7792 GCTGGCGCTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAATCGAGCTCAAG 7851
QY 4114 TCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGCGTTTCCCGCTCGAAGCTC 4173
DB 7852 TCAGAGTGGCGAAACCCGACAGGACTATAAAGATACGAGCGTTTCCCGCTCGAAGCTC 7911
QY 4174 CTTGTCGCTCTCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCTTCTCC 4233
DB 7912 CTTGTCGCTCTCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCTTCTCC 7971
QY 4234 TCCGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGAGGTATCTCAGTTTCGCTGAGGT 4293
DB 7972 TCCGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGAGGTATCTCAGTTTCGCTGAGGT 8031
QY 4294 CGTTTCGCTCCAAGTGGGCTGTGACAGAAACCCCGTTTCCGCGGACCGCTGGCGCTT 4353
DB 8032 CGTTTCGCTCCAAGTGGGCTGTGACAGAAACCCCGTTTCCGCGGACCGCTGGCGCTT 8091
QY 4354 ATCCGGTAACTATGCTTCCAGTCCAAACCCGGTAAAGACAGCTTATCGCACTGGCAGC 4413
DB 8092 ATCCGGTAACTATGCTTCCAGTCCAAACCCGGTAAAGACAGCTTATCGCACTGGCAGC 8151
QY 4414 AGCCACTGGTAAACAGGATTAAGCAGGAGGTATGAGGCGGTGCTACAGAGTTCTTGAA 4473
DB 8152 AGCCACTGGTAAACAGGATTAAGCAGGAGGTATGAGGCGGTGCTACAGAGTTCTTGAA 8211
QY 4474 GTGTTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA 4533
DB 8212 GTGTTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAA 8271
QY 4534 GCCAGTTACCTTCGGAAGAGTGTGAGTCTTGTATCGGCAAAACAAACCCAGCTGG 4593
DB 8272 GCCAGTTACCTTCGGAAGAGTGTGAGTCTTGTATCGGCAAAACAAACCCAGCTGG 8331
QY 4594 TAGCGGTGGTTTTTGTGTAAGGAGGAGTATGCGGAGGAAAAAAGGATCTCAAGA 4653
DB 8332 TAGCGGTGGTTTTTGTGTAAGGAGGAGTATGCGGAGGAAAAAAGGATCTCAAGA 8391

QY 4654 AGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGCTTAAG 4713
DB 8392 AGATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAAACTCAGCTTAAG 8451
QY 4714 GATTTTGGTCAAGAGATTAATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATG 4773
DB 8452 GATTTTGGTCAAGAGATTAATCAAAAAGGATCTTCACTAGATCCTTTTAAATTTAAATG 8511
QY 4774 AAGTTTTAAATCAATCTAAAGTATATATGATTAATCTTGGTCTGACAGTTTACCAATGCTT 4833
DB 8512 AAGTTTTAAATCAATCTAAAGTATATATGATTAATCTTGGTCTGACAGTTTACCAATGCTT 8571
QY 4834 AATCAGTAGGCACTTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTTGCCTGACT 4893
DB 8572 AATCAGTAGGCACTTATCTCAGCGATCTGTCTATTTTGGTTCATCATAGTTTGCCTGACT 8631
QY 4894 CCCCCTGCTGATAGTAATCTACGATACGGAGGGCTTACCATCTGGGCCCGAGTCTGCAAT 4953
DB 8632 CCCCCTGCTGATAGTAATCTACGATACGGAGGGCTTACCATCTGGGCCCGAGTCTGCAAT 8691
QY 4954 GATACCGCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCAGCGG 5013
DB 8692 GATACCGCAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCCAGCAGCGG 8751
QY 5014 AAGGGCCGAGCGCAGAGTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTG 5073
DB 8752 AAGGGCCGAGCGCAGAGTGGTCTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAATTG 8811
QY 5074 TTGCCGGGAAGCTAGAGTAAGTAGTTGGCAGTTAATAGTTTGGCGCAACGTTTGGCCAT 5133
DB 8812 TTGCCGGGAAGCTAGAGTAAGTAGTTGGCAGTTAATAGTTTGGCGCAACGTTTGGCCAT 8871
QY 5134 TGCTACAGCATCTGCTGCTCAGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGCTTC 5193
DB 8872 TGCTACAGCATCTGCTGCTCAGCTCGCTGTTTGGTATGGCTTCAATTCAGCTCCGCTTC 8931
QY 5194 CCAACCATCAAGCGAGTTTACATGATCCCGCCATGTTTGTGCAAAAAGCGTTAGCTCCTT 5253
DB 8932 CCAACCATCAAGCGAGTTTACATGATCCCGCCATGTTTGTGCAAAAAGCGTTAGCTCCTT 8991
QY 5254 CGGTCTCCGATGCTGTTGTCAGAAAGTAAAGTTGGCGCAGTGTATATCATGTTTATGGC 5313
DB 8992 CGGTCTCCGATGCTGTTGTCAGAAAGTAAAGTTGGCGCAGTGTATATCATGTTTATGGC 9051
QY 5314 AGCACTGCAATATCTTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGA 5373
DB 9052 AGCACTGCAATATCTTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGA 9111
QY 5374 GTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGGC 5433
DB 9112 GTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGGC 9171
QY 5434 GTCAATAACGGGATAATAACCGCGCCACATAGCAGAACTTTTAAAGTCTCATCATTTGAAA 5493
DB 9172 GTCAATAACGGGATAATAACCGCGCCACATAGCAGAACTTTTAAAGTCTCATCATTTGAAA 9231
QY 5494 AGTTCTTCCGGGCGGAAACTCTCAGGATCTTACCGCTGTTGAGATCAGTTTCGATGTA 5553
DB 9232 AGTTCTTCCGGGCGGAAACTCTCAGGATCTTACCGCTGTTGAGATCAGTTTCGATGTA 9291
QY 5554 ACCCACTCCGTCACCCCACTGATCTTTCAGCATCTTTTACCTTTTCCAGCGGTTTCTGGGTG 5613
DB 9292 ACCCACTCCGTCACCCCACTGATCTTTCAGCATCTTTTACCTTTTCCAGCGGTTTCTGGGTG 9351
QY 5614 AGCAAAAACAGGAAGCAAAATCCGCGCAAAAAGGGAATAAGGGCGACACGGAATTTGTTG 5673
DB 9352 AGCAAAAACAGGAAGCAAAATCCGCGCAAAAAGGGAATAAGGGCGACACGGAATTTGTTG 9411
QY 5674 AATACTCATCTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTTATGTTCTCAT 5733
DB 9412 AATACTCATCTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTTATGTTCTCAT 9471
QY 5734 GAGCGGATACATATTTGTAATGTTATTTAGAAAAATAAACAATAGGGGTTCCCGGCACATT 5793

Db	9472	GAGCGGATACATATTGGAATGTATTAGAAAAATAAACAAATAGGGGTTCCGGGCACATT	9531
Qy	5794	TCCCCGAAAAAGTCCACCTGACGCTTAAGAAACCATTTATCATGACATTAACTATATA	5853
Db	9532	TCCCCGAAAAAGTCCACCTGACGCTTAAGAAACCATTTATCATGACATTAACTATATA	9591
Qy	5854	AAATAGGCGTATCACGAGGCCCTTTCGTC	5882
Db	9592	AAATAGGCGTATCACGAGGCCCTTTCGTC	9620

Search completed: April 30, 2005, 05:45:05
Job time : 24242 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2005, 18:19:56 ; Search time 2706 Seconds
(without alignments)

12867.671 Million cell updates/sec

Title: US-10-715-665-6

Perfect score: 5882

Sequence: 1 tcgcgctttcgtgatgac.....tatcagaggcccttcctgc 5882

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5882	100.0	5882	5	AAD06793
2	4617.2	78.5	5459	5	AAD06792
3	4190	71.2	5128	5	AAD06791
4	3713.6	63.1	6299	4	Aaf83669
5	2658.4	45.2	4276	5	AAD06790
6	2654.4	45.1	10263	12	AD104102
7	2601.2	44.2	4282	4	Aaf83668
8	2419.2	41.1	6236	10	ACC71561
9	2399.4	40.8	6050	3	AAA47797
10	2397.8	40.8	6050	12	ADK15555
11	2395.6	40.7	6148	10	ACC71544
12	2363	40.2	6258	10	ACC71554
13	2333.2	39.7	7551	12	AD172981
14	2321.4	39.5	7626	10	AD864206
15	2290.6	38.9	6577	6	ABK91498
16	2283.4	38.8	6505	6	ABK91463
17	2283	38.8	6460	6	ABK91482
18	2281.8	38.8	9620	4	Aaf83667
19	2281.8	38.8	9620	4	Aaf83666
20	2280.4	38.8	6577	6	ABK91499

21	2280.4	38.8	7608	10	ADE64205
22	2275.8	38.7	6438	6	ABK91605
23	2275.8	38.7	6577	6	ABK91480
24	2270.8	38.6	6473	6	ABK91548
25	2269.2	38.6	6607	6	ABK91456
26	2266	38.5	6486	6	ABK91606
27	2265.4	38.5	8575	2	AAQ40324
28	2263.6	38.5	6605	6	ABK91553
29	2263.6	38.5	6605	6	ABK91554
30	2262	38.5	6538	6	ABK91481
31	2261.6	38.4	6608	6	ABK91570
32	2260.6	38.4	7178	12	ADL72979
33	2260	38.4	6602	6	ABK91551
34	2259.8	38.4	6609	6	ABK91532
35	2259.6	38.4	6536	6	ABK91555
36	2256.6	38.4	6505	6	ABK91464
37	2256.6	38.4	6565	3	ABZ45925
38	2256	38.4	6545	6	ABK91563
39	2255.4	38.3	8911	12	ADN07735
40	2252	38.3	6561	10	ACC71524
41	2252	38.3	6577	6	ABK91592
42	2251.4	38.3	6466	6	ABK91578
43	2251.4	38.3	6538	6	ABK91593
44	2251.4	38.3	6586	6	ABK91576
45	2250.8	38.3	6614	6	ABK91562

ALIGNMENTS

RESULT 1
AAD06793

ID AAD06793 standard; DNA; 5882 BP.

AC AAD06793;

DT 06-AUG-2001 (first entry)

DE Plasmid pCMV-II-E2661-SAG.

XX Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;
KW HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;
XX plasmid pCMV-II-E2661-SAG; cyclic; circular; ds.

OS Human cytomegalovirus.

OS Homo sapiens.

OS Hepatitis C virus.

OS Hepatitis B virus.

OS Bos sp.

OS Unidentified.

OS Chimeric.

XX Key Location/Qualifiers
CDS 1992..3587

FT /tag= a
FT /product= "HCV 661 E2 envelope glycoprotein-HBsAg S
FT domain"

FT misc_feature 1992..2900

FT /tag= b
FT /note= "661 E2 coding sequence from HCV"

FT misc_feature 2907..3587

FT /tag= c
FT /note= "HBsAg S domain coding sequence from HBV"

XX WO200138358-A2.

PN 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032249.

XX 24-NOV-1999; 99US-0167224P.

XX (CHIR) CHIRON CORP.

XX	Selby M, Glazer E, Houghton M;	
PI	WPI; 2001-367661/38.	
XX	P-PSDB; AAE02622.	
DR		
XX	Virus-like particle for use as an immunogen, comprising a first hepatitis B virus surface antigen (HBsAg) and chimeric antigen comprising a second HBsAg covalently linked to hepatitis C immunogenic polypeptide.	
PT	Claim 37; Fig 4; 115pp; English.	
PT	The invention relates to a virus-like particle for use as an immunogen, comprising a first hepatitis B virus surface antigen (HBsAg) and a chimeric antigen comprising a second HBsAg which is covalently linked to hepatitis C virus (HCV) immunogenic polypeptide, where the first and the second HBsAg each comprise a substantially complete S domain. The virus-like particle is useful as immunogen and as vaccine. The present sequence is plasmid pCMV-II-E2661-sAg containing coding sequences for 661 E2 envelope glycoprotein of HCV and S domain of HBsAg inserted into pCMVII. The plasmid pCMVII is a pUC19-based cloning vector designed for expression in mammalian cells. It comprises human cytomegalovirus (CMV) immediate early (IE) enhancer/promoter, human CMV intron A, a human tissue plasminogen activator (tPA) leader, a bovine growth hormone poly A terminator (BGHt), a ColEI origin of replication and an ampicillin resistance (Amp R) gene. The pCMV-II-E2661-sAg plasmid is used for production of virus-like particle of the invention	
XX	Query Match 100.0%; Score 5882; DB 5; Length 5882; Best Local Similarity 100.0%; Pred. No. 0; Matches 5882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
SQ	Sequence 5882 BP; 1378 A; 1543 C; 1469 G; 1492 T; 0 U; 0 Other;	
QY	1 TC CGCGGTTTCGGTGATGACGGTGA AAAACCTCTGACACATGACGTCCCGGAGACGGTCA 60	601 GGAGCTTTCATTTGAGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCCATTTGGCAGTA 660
DB	1 TC CGCGGTTTCGGTGATGACGGTGA AAAACCTCTGACACATGACGTCCCGGAGACGGTCA 60	601 GGAGCTTTCATTTGAGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCCATTTGGCAGTA 660
QY	61 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAGGCGGGTG 120	661 CATCAAGTGTATCATATGCAAGTCCGCGCCCTATTGACGTCAATGACGGTAAATGGCCC 720
DB	61 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAGGCGGGTG 120	661 CATCAAGTGTATCATATGCAAGTCCGCGCCCTATTGACGTCAATGACGGTAAATGGCCC 720
QY	121 TTGGCGGGTGTGGGGCTGGCTTAACCTATGCGGCATCAGACAGATGCTGAGAGTGC 180	721 GCCTGGCATTTAGCCCCAGTACATGACCTTTACGGGACTTTTCTTACTTTGGCAGTACATCTAC 780
DB	121 TTGGCGGGTGTGGGGCTGGCTTAACCTATGCGGCATCAGACAGATGCTGAGAGTGC 180	721 GCCTGGCATTTAGCCCCAGTACATGACCTTTACGGGACTTTTCTTACTTTGGCAGTACATCTAC 780
QY	181 ACCATATGAAGCTTTTTC AAAAGCCTTAGGCCCTCCAAAAGCCTCTCACTACTTCTGG 240	781 GTATTAGTCATCGCTATTACATGATGCGGTTTGGCAGTACACCAATGGCGGTGGA 840
DB	181 ACCATATGAAGCTTTTTC AAAAGCCTTAGGCCCTCCAAAAGCCTCTCACTACTTCTGG 240	781 GTATTAGTCATCGCTATTACATGATGCGGTTTGGCAGTACACCAATGGCGGTGGA 840
QY	241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGTCATAATAAAAAAATAGTCAGCCA 300	841 TAGCGGTTTGACTCAGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900
DB	241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGTCATAATAAAAAAATAGTCAGCCA 300	841 TAGCGGTTTGACTCAGCGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900
QY	301 TGGGGCGGAGAAATGGCGGAACCTGGGCGGAGGGAATTTGGCTATTGGCCATTGCAAT 360	901 TTTTGGCACA AAAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCTTTGACG 960
DB	301 TGGGGCGGAGAAATGGCGGAACCTGGGCGGAGGGAATTTGGCTATTGGCCATTGCAAT 360	901 TTTTGGCACA AAAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCTTTGACG 960
QY	361 ACCTTGATCTATCATATATGATACATTTATTTGGCTCATGTCCAATATGACCGCCA 420	961 CAAATGGGCGGTAGGCGTGTACGGTGGAGGTCTATATAGCAGAGCTGTTTTAGTAGAAC 1020
DB	361 ACCTTGATCTATCATATATGATACATTTATTTGGCTCATGTCCAATATGACCGCCA 420	961 CAAATGGGCGGTAGGCGTGTACGGTGGAGGTCTATATAGCAGAGCTGTTTTAGTAGAAC 1020
QY	421 TGTGTGACATTGATTATGACTAGTTAATAGTAATCAATTAACGGGGTCATTAGTTCAAT 480	1021 CGTCAGATCGCTGGAGAGCCCATCCACGCTGTTTTGACCTCCATAGAGACACCCGGAC 1080
DB	421 TGTGTGACATTGATTATGACTAGTTAATAGTAATCAATTAACGGGGTCATTAGTTCAAT 480	1021 CGTCAGATCGCTGGAGAGCCCATCCACGCTGTTTTGACCTCCATAGAGACACCCGGAC 1080
QY	481 AGCCCATATATGAGTTCGCGTTACATTAACCTTACGGTAAATGGCGGCTGGCTGACCG 540	1081 CGATCCAGGCTTCGCGCGCGGGAACCGGTGCATTTGGAAACCGGATTTCCCGTGCACAGAGT 1140
DB	481 AGCCCATATATGAGTTCGCGTTACATTAACCTTACGGTAAATGGCGGCTGGCTGACCG 540	1081 CGATCCAGGCTTCGCGCGCGGGAACCGGTGCATTTGGAAACCGGATTTCCCGTGCACAGAGT 1140
QY	541 CCCAAGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600	1141 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1200
DB	541 CCCAAGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600	1141 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1200

Qy	1681	AGTGTGCGGCACAAAGCCGTGGCGGTAGGGTATGTGTCTGAAATGAGCTCGAGATTGG	1740
Db	1681	AGTGTGCGGCACAAAGCCGTGGCGGTAGGGTATGTGTCTGAAATGAGCTCGAGATTGG	1740
Qy	1741	GCTCGCACTTGGACGAGATGGAAGACTTAAAGCAGCGCGCAGAAAGATGCGAGCAGCT	1800
Db	1741	GCTCGCACTTGGACGAGATGGAAGACTTAAAGCAGCGCGCAGAAAGATGCGAGCAGCT	1800
Qy	1801	GAGTGTGTATTTCTGATAAGAGTCAGAGGTAACTCCCGTTGCGGTGCTGTAAACGCTGG	1860
Db	1801	GAGTGTGTATTTCTGATAAGAGTCAGAGGTAACTCCCGTTGCGGTGCTGTAAACGCTGG	1860
Qy	1861	AGGCACTGTAGTCTGAGCAGTACTCTGTCTGCGCGCGCGCCACAGACATAATAGCT	1920
Db	1861	AGGCACTGTAGTCTGAGCAGTACTCTGTCTGCGCGCGCGCCACAGACATAATAGCT	1920
Qy	1921	GACAGACTACAGACTGTTCTTTTCCATGGGTCTTTTCTGCACTCACCCTCGTCGAGAA	1980
Db	1921	GACAGACTACAGACTGTTCTTTTCCATGGGTCTTTTCTGCACTCACCCTCGTCGAGAA	1980
Qy	1981	TTCAAGCAATCATGGATGCAATGAAGAGAGGCTCTCTGTCTGCTGCTGTGTGGAG	2040
Db	1981	TTCAAGCAATCATGGATGCAATGAAGAGAGGCTCTCTGTCTGCTGCTGTGTGGAG	2040
Qy	2041	CAGTCTTTCGTTTCGCCAGCGCTAGCGAAACCCACGTCACCGGGGGAAGTCCCGGCCACA	2100
Db	2041	CAGTCTTTCGTTTCGCCAGCGCTAGCGAAACCCACGTCACCGGGGGAAGTCCCGGCCACA	2100
Qy	2101	CTGTGCTGATTTGTTAGCTCTCTCGACAGCGGCTCTGCTGTCTGCTGCTGTGTGGAG	2160
Db	2101	CTGTGCTGATTTGTTAGCTCTCTCGACAGCGGCTCTGCTGTCTGCTGCTGTGTGGAG	2160
Qy	2161	ACACCAACGCGAGTTGCGACTCAATAGCAGCGCCCTGAACTGCAATGATAGCTCAACA	2220
Db	2161	ACACCAACGCGAGTTGCGACTCAATAGCAGCGCCCTGAACTGCAATGATAGCTCAACA	2220
Qy	2221	CCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAAGGCTGTCTTGAGA	2280
Db	2221	CCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAAGGCTGTCTTGAGA	2280
Qy	2281	GGCTAGCAGCTGCGGACCCCTTACCGATTTTGAACGAGGCTGGGGCCCTATCAGTTATG	2340
Db	2281	GGCTAGCAGCTGCGGACCCCTTACCGATTTTGAACGAGGCTGGGGCCCTATCAGTTATG	2340
Qy	2341	CCAAAGGAAGCGGCCCGACAGCGCCCTACTGTCTGCGCACTACCCCCCAAAAACCTTGGC	2400
Db	2341	CCAAAGGAAGCGGCCCGACAGCGCCCTACTGTCTGCGCACTACCCCCCAAAAACCTTGGC	2400
Qy	2401	GTATTTGCGCGGAAGAGTGTGTGTGTCGGTATATTGCTTCACTCCAGCGCCCGTGG	2460
Db	2401	GTATTTGCGCGGAAGAGTGTGTGTGTCGGTATATTGCTTCACTCCAGCGCCCGTGG	2460
Qy	2461	TGTTGGGAACGACCGACAGTTCGGCGCGCCGACCTACAGCTGGGGTGAATAACGG	2520
Db	2461	TGTTGGGAACGACCGACAGTTCGGCGCGCCGACCTACAGCTGGGGTGAATAACGG	2520
Qy	2521	ACGTTCTTCTGCTTAAACAATACAGGCGACCGCTGGGCAATTGGTTCGGTTGACCTTGG	2580
Db	2521	ACGTTCTTCTGCTTAAACAATACAGGCGACCGCTGGGCAATTGGTTCGGTTGACCTTGG	2580
Qy	2581	TGAACCTCAACTGGATTTCACCAAAAGTGTGCGAGCGCTCTCTGTGTCTATCGAGGGGCGG	2640
Db	2581	TGAACCTCAACTGGATTTCACCAAAAGTGTGCGAGCGCTCTCTGTGTCTATCGAGGGGCGG	2640
Qy	2641	GCAACAAACACCTTGCACTGCCCACTGATTTGCTTTCGCAAGCATTCGGGACGCCACATCT	2700
Db	2641	GCAACAAACACCTTGCACTGCCCACTGATTTGCTTTCGCAAGCATTCGGGACGCCACATCT	2700
Qy	2701	CTCGGTGCGGCTCCGGTCCCTGATACACCCAGGTGCTGTGCTACTCCGCTATAGGC	2760
Db	2701	CTCGGTGCGGCTCCGGTCCCTGATACACCCAGGTGCTGTGCTACTCCGCTATAGGC	2760
Qy	2761	TTTGGCATATTCTTGTATCCATCAACTACACCATATTATAAAATCAGGATGTACGTGGAG	2820

Db	2761	TTTGGCATATTCTTGTATCCATCAACTACACCATATTATAAAATCAGGATGTACGTGGAG	2820
Qy	2821	GGGTCCGAACACAGGCTTGGAGCTGCTGCAACTGGAACGCGCGGGGCAACGTTGGATCTGG	2880
Db	2821	GGGTCCGAACACAGGCTTGGAGCTGCTGCAACTGGAACGCGCGGGGCAACGTTGGATCTGG	2880
Qy	2881	AAGATAGGGAACAGGCTCGAGATCGATATGGAGAACATCATCAGGATTTCTTAGGACCCC	2940
Db	2881	AAGATAGGGAACAGGCTCGAGATCGATATGGAGAACATCATCAGGATTTCTTAGGACCCC	2940
Qy	2941	TGCTCGTGTTAACAGGCGGGTCTTTCTTGTGTGAACAAGATCTCTCAATACCGCAGAGTC	3000
Db	2941	TGCTCGTGTTAACAGGCGGGTCTTTCTTGTGTGAACAAGATCTCTCAATACCGCAGAGTC	3000
Qy	3001	TAGACTCGTGTGAGTCTCTCAATTTTCTAGGGGATCTCCGCTGTGTCTTGGCCAAA	3060
Db	3001	TAGACTCGTGTGAGTCTCTCAATTTTCTAGGGGATCTCCGCTGTGTCTTGGCCAAA	3060
Qy	3061	ATTTCGAGTCCCCAACCTCCAATCACTCACAACTCTCTCTCATCTCTGCTCTGCTGCT	3120
Db	3061	ATTTCGAGTCCCCAACCTCCAATCACTCACAACTCTCTCTCATCTCTGCTCTGCTGCT	3120
Qy	3121	ATCGCTGAGTGTCTGCGGCTTTTATCATATTTCTCTTTCATCTCTGCTCTGCTGCTCA	3180
Db	3121	ATCGCTGAGTGTCTGCGGCTTTTATCATATTTCTCTTTCATCTCTGCTCTGCTGCTCA	3180
Qy	3181	TCCTCTTATTTGGTCTTCTGGAATTAACAAGTATGTGCGGCTTGTCTCTAATTTCCAG	3240
Db	3181	TCCTCTTATTTGGTCTTCTGGAATTAACAAGTATGTGCGGCTTGTCTCTAATTTCCAG	3240
Qy	3241	GATCAACAACAACAGTACGGGACCATGCAAAACCTGCAAGCTCTCTGCTCAAGGCAACT	3300
Db	3241	GATCAACAACAACAGTACGGGACCATGCAAAACCTGCAAGCTCTCTGCTCAAGGCAACT	3300
Qy	3301	CTATGTTTCCCTCATGTTGCTGTACAAAACCTTAGGATGGAATTTGCACTGTATTTCCCA	3360
Db	3301	CTATGTTTCCCTCATGTTGCTGTACAAAACCTTAGGATGGAATTTGCACTGTATTTCCCA	3360
Qy	3361	TCCCATGCTCTGGGCTTTTCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTCTCTT	3420
Db	3361	TCCCATGCTCTGGGCTTTTCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTCTCTT	3420
Qy	3421	GGCTCAGTTTACTAGTGCCTATTTGTTTTCAGTGGTTCGTAGGGCTTTTCCCCACCTGTTGGC	3480
Db	3421	GGCTCAGTTTACTAGTGCCTATTTGTTTTCAGTGGTTCGTAGGGCTTTTCCCCACCTGTTGGC	3480
Qy	3481	TTTCAGCTATATGGATGATGTGTATTTGGGGCCAAAGTCTGTACAGCATCGTGAAGTCCCT	3540
Db	3481	TTTCAGCTATATGGATGATGTGTATTTGGGGCCAAAGTCTGTACAGCATCGTGAAGTCCCT	3540
Qy	3541	TTATACCGCTGTATCAAAATTTCTTTTGTCTCTGGGTATACATTTAAGAATTCAGACTCG	3600
Db	3541	TTATACCGCTGTATCAAAATTTCTTTTGTCTCTGGGTATACATTTAAGAATTCAGACTCG	3600
Qy	3601	AGCAAGCTTAGAAAGCGCGCCAAAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA	3660
Db	3601	AGCAAGCTTAGAAAGCGCGCCAAAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA	3660
Qy	3661	TCAGCTTCGACTGTGCTTCTAGTTCGCAAGCATCTGTGTTGTCGCTCTCCCGCTGCTCT	3720
Db	3661	TCAGCTTCGACTGTGCTTCTAGTTCGCAAGCATCTGTGTTGTCGCTCTCCCGCTGCTCT	3720
Qy	3721	TCCTTCGACCTGGAAGGTGCCACTTCTATTTCTGGGGGTGGGTGGGGCAGGACAGCAAG	3780
Db	3721	TCCTTCGACCTGGAAGGTGCCACTTCTATTTCTGGGGGTGGGTGGGGCAGGACAGCAAG	3780
Qy	3781	TCGCATTGTCTAGTAGTGTCTCTTCTATTTCTGGGGGTGGGTGGGGCAGGACAGCAAG	3840
Db	3781	TCGCATTGTCTAGTAGTGTCTCTTCTATTTCTGGGGGTGGGTGGGGCAGGACAGCAAG	3840
Qy	3841	GGGGAGATTGGGAAGCAATACAGGCAATGTGGGAGCTTTCCGCTTCTCTGCTCTCAC	3900

Db 3841 GGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGAGCTCTTCGCTTCCTCGCTCAC 3900
QY 3901 TGA CTGCTCGCTCGGTCTGCTTGGCTGCGGCGAGCGGTATACGTCACTCAAAGCGGT 3960
Db 3901 TGACTCGCTCGCTCGGTCTGCGGTGCGGCGAGCGGTATCAGCTCACTCAAAGCGGT 3960
QY 3961 AATACGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAAATGTGAGCAAAAGGCCA 4020
Db 3961 AATACGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAAATGTGAGCAAAAGGCCA 4020
QY 4021 GCAAAAGGCCAGGAAACCGTAAAAAGCCGCTGGCTGGCGTTTTTCCATAGGCTCGGCC 4080
Db 4021 GCAAAAGGCCAGGAAACCGTAAAAAGCCGCTGGCTGGCGTTTTTCCATAGGCTCGGCC 4080
QY 4081 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGACT 4140
Db 4081 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGACT 4140
QY 4141 ATAAAGATACCGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTT 4200
Db 4141 ATAAAGATACCGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTT 4200
QY 4201 GCCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCAATG 4260
Db 4201 GCCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCAATG 4260
QY 4261 CTCACGCTGATGATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCA 4320
Db 4261 CTCACGCTGATGATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCA 4320
QY 4321 CGAACCCCGTTTCAGCCGACCGCTCGGCTTATCGGTAACTATCGTTGAGTCCAA 4380
Db 4321 CGAACCCCGTTTCAGCCGACCGCTCGGCTTATCGGTAACTATCGTTGAGTCCAA 4380
QY 4381 CCCGCTAAGACACGACTTATCGGCACCTGGCAGCAGCACCTGGTAAACAGGATTAGCAGAGC 4440
Db 4381 CCCGCTAAGACACGACTTATCGGCACCTGGCAGCAGCACCTGGTAAACAGGATTAGCAGAGC 4440
QY 4441 GAGGTATGATAGGCGGTCTACAGATTCTTGAAGTGGTGCCCTAACTACGGCTACACTAG 4500
Db 4441 GAGGTATGATAGGCGGTCTACAGATTCTTGAAGTGGTGCCCTAACTACGGCTACACTAG 4500
QY 4501 AAGGACAGATTATGGTATCTGCGCTCTGCTGAGCCAGTACCTTCGGAAAAAGAGTTGG 4560
Db 4501 AAGGACAGATTATGGTATCTGCGCTCTGCTGAGCCAGTACCTTCGGAAAAAGAGTTGG 4560
QY 4561 TAGCTCTTGATCCGGCAAAACAAACCCACCGCTGCTAGCGGTGGTTTTTTTTTTGTTGCAAGCA 4620
Db 4561 TAGCTCTTGATCCGGCAAAACAAACCCACCGCTGCTAGCGGTGGTTTTTTTTTTGTTGCAAGCA 4620
QY 4621 GCAGATTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCTTTTCTACGGGGTC 4680
Db 4621 GCAGATTACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCTTTTCTACGGGGTC 4680
QY 4681 TGACGCTCAGTGGAACGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAAG 4740
Db 4681 TGACGCTCAGTGGAACGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAAG 4740
QY 4741 GATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATA 4800
Db 4741 GATCTTCACTAGATCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATA 4800
QY 4801 TGAGTAAACTTGTCTGACAGTTACCAATGCTTTAAATCAGTGAGGCACTTATCTCAGCGAT 4860
Db 4801 TGAGTAAACTTGTCTGACAGTTACCAATGCTTTAAATCAGTGAGGCACTTATCTCAGCGAT 4860
QY 4861 CTGCTATTTTCTGTTTCACTAGTTCCTGACTCCCGCTCGTGTAGTAACCTACGATACG 4920
Db 4861 CTGCTATTTTCTGTTTCACTAGTTCCTGACTCCCGCTCGTGTAGTAACCTACGATACG 4920
QY 4921 GGAGGGCTTACCACTCTGGCCCCCAGTGTGCAATGATACCGCAGAGACCCAGCTCACC 4980
Db 4921 GGAGGGCTTACCACTCTGGCCCCCAGTGTGCAATGATACCGCAGAGACCCAGCTCACC 4980

QY 4981 TCCAGATTATTCAGCAATAAACACAGCCAGCGCGAAGCGGCGAGAGTGGTCTCTGC 5040
Db 4981 TCCAGATTATTCAGCAATAAACACAGCCAGCGCGAAGCGGCGAGAGTGGTCTCTGC 5040
QY 5041 AACTTTATCCGCTCCATCCAGTCTAATAATTTGTTCCCGGGAAGCTAGAGTAAGTAGTTTC 5100
Db 5041 AACTTTATCCGCTCCATCCAGTCTAATAATTTGTTCCCGGGAAGCTAGAGTAAGTAGTTTC 5100
QY 5101 GCCAGTTAATAGTTTTCGCAACGTTTTCGCAATTCGTACAGGCATCGTGGTGTACGCTC 5160
Db 5101 GCCAGTTAATAGTTTTCGCAACGTTTTCGCAATTCGTACAGGCATCGTGGTGTACGCTC 5160
QY 5161 GTCGTTTGTATGCTTCATTCAGCTCCCGTTCCTCCAAAGATCAAGGCGAGTTACATGATC 5220
Db 5161 GTCGTTTGTATGCTTCATTCAGCTCCCGTTCCTCCAAAGATCAAGGCGAGTTACATGATC 5220
QY 5221 CCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGCTCCTCCGATCGTTGTTCAGAAAGTAA 5280
Db 5221 CCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGCTCCTCCGATCGTTGTTCAGAAAGTAA 5280
QY 5281 GTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTGCAT 5340
Db 5281 GTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTGCAT 5340
QY 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAAATA 5400
Db 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAAATA 5400
QY 5401 GTGTATGCGCGCAGCCGAGTTGCTCTTTCGCGCTCAATACGGGATAATACCGCGCCACA 5460
Db 5401 GTGTATGCGCGCAGCCGAGTTGCTCTTTCGCGCTCAATACGGGATAATACCGCGCCACA 5460
QY 5461 TAGCAGAACTTTTAAAGTCTCATCATTTGGAACCGTTCTTCGGGCGGAAAACTCTCAAG 5520
Db 5461 TAGCAGAACTTTTAAAGTCTCATCATTTGGAACCGTTCTTCGGGCGGAAAACTCTCAAG 5520
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGACCCCACTGTCTTC 5580
Db 5521 GATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCCACTCGTGACCCCACTGTCTTC 5580
QY 5581 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTTGAGCAAAAAACAGGAAGCGCAAAATGCCGC 5640
Db 5581 AGCATCTTTTACTTTTACCAGCGTTTCTCGGTTGAGCAAAAAACAGGAAGCGCAAAATGCCGC 5640
QY 5641 AAAAAAGGGAATAAGGCGGACACGGAATGTTGAATFACTATACTCTTCTCTTTTCAATA 5700
Db 5641 AAAAAAGGGAATAAGGCGGACACGGAATGTTGAATFACTATACTCTTCTCTTTTCAATA 5700
QY 5701 TTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
Db 5701 TTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
QY 5761 GAAAAATAACAAATAGGGGTTCCGCGCACATTTTCCCGAAAGTGCCACCTCAGCTCTA 5820
Db 5761 GAAAAATAACAAATAGGGGTTCCGCGCACATTTTCCCGAAAGTGCCACCTCAGCTCTA 5820
QY 5821 AGAAACCATTTATTCATGACATTAACCTTATAAAAAATAGCGGTATCACGAGGCCCTTCG 5880
Db 5821 AGAAACCATTTATTCATGACATTAACCTTATAAAAAATAGCGGTATCACGAGGCCCTTCG 5880
QY 5881 TC 5882
Db 5881 TC 5882
RESULT 2
AAD06792
ID AAD06792 standard; DNA; 5459 BP.
XX
AAD06792;
XX
06-AUG-2001 (first entry)
DT

```

XX DE Plasmid pCMVII opti 330 E1/SAG.
XX KW Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;
XX KW HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;
XX KW plasmid pCMVII opti 330 E1/SAG; cyclic; circular; ds.
XX OS Human cytomegalovirus.
OS OS Homo sapiens.
OS OS Hepatitis C virus.
OS OS Hepatitis B virus.
OS OS Boe sp.
OS OS Unidentified.
OS OS Chimeric.
XX PH Key Location/Qualifiers
XX FT CDS 1992..3164
XX FT /*tag= a
XX FT /product= "HCV 330 E1 envelope glycoprotein-HBsAg S
XX FT domain"
XX FT misc_feature 1992..2483
XX FT /*tag= b
XX FT /note= "330 E1 coding sequence from HCV"
XX FT misc_feature 2484..3164
XX FT /*tag= c
XX FT /note= "S domain coding sequence from HBsAg"
XX FT WO200138358-A2.
XX PN
XX XX
XX PD 31-MAY-2001.
XX XX
XX PF 22-NOV-2000; 2000WO-US032249.
XX XX
XX PR 24-NOV-1999; 99US-0167224P.
XX XX
XX PA (CHIR ) CHIRON CORP.
XX XX
XX PI Selby M, Glazer E, Houghton M;
XX XX
XX DR WPI; 2001-367661/38.
XX DR P-PSDB; AAE02621.
XX XX
XX PT Virus-like particle for use as an immunogen, comprising a first hepatitis
XX PT B virus surface antigen (HBsAg) and chimeric antigen comprising a second
XX PT HBsAg covalently linked to hepatitis C immunogenic polypeptide.
XX XX
XX PS Claim 33; Fig 3; 115pp; English.
XX XX
XX CC The invention relates to a virus-like particle for use as an immunogen,
XX CC comprising a first hepatitis B virus surface antigen (HBsAg) and a
XX CC chimeric antigen comprising a second HBsAg which is covalently linked to
XX CC an hepatitis C virus (HCV) immunogenic polypeptide, where the first and
XX CC the second HBsAg each comprise a substantially complete S domain. The
XX CC virus-like particle is useful as immunogen and as vaccine. The present
XX CC sequence is plasmid pCMVII opti 330 E1/SAG containing coding sequences
XX CC for 330 E1 envelope glycoprotein of HCV and S domain of HBsAg inserted
XX CC into pCMVII. The plasmid pCMVII is a pUC19-based cloning vector designed
XX CC for expression in mammalian cells. It comprises human cytomegalovirus
XX CC (CMV) immediate early (IE) enhancer/promoter, human CMV intron A, a human
XX CC tissue plasminogen activator (tPA) leader, a bovine growth hormone poly A
XX CC terminator (BGHT), a ColEI origin of replication and an ampicillin
XX CC resistance (Amp R) gene. The pCMVII opti 330 E1/SAG plasmid is used for
XX CC production of virus-like particle of the invention
XX XX
XX SQ Sequence 5459 BP; 1267 A; 1458 C; 1357 G; 1377 T; 0 U; 0 Other;
XX
XX Query Match 78.5%; Score 4617.2; DB 5; Length 5459;
XX Best Local Similarity 89.1%; Pred. No. 0;
XX Matches 5241; Conservative 0; Mismatches 218; Indels 423; Gaps 7;
XX
XX 1 TCGGCGCTTCGGTGATGACGGTGAAAACCTCTGACACATGCGAGCTCCCGAGACGGTCA 60
XX Db 1 TCGGCGCTTCGGTGATGACGGTGAAAACCTCTGACACATGCGAGCTCCCGAGACGGTCA 60

```

QY 1141 GAGTAAAGTACCGCTATAGACTCTATAGGCACACCCCTTGGCTCTTATGATGCTATA 1200
DB 1141 GAGCTAAGTACCGCCCTATAGACTCTATAGGCACACCCCTTGGCTCTTATGATGCTATA 1200
QY 1201 CTGTTTGTGGCTTGGGSCCTATACACCCCGCTCCCTTATAGCTATAGGTATAGCT 1260
DB 1201 CTGTTTGTGGCTTGGGSCCTATACACCCCGCTCCCTTATAGCTATAGGTATAGCT 1260
QY 1261 TAGCCTATAGGTGTGGTATTACCACTATTAGCACTCCCTATTGAGTATAGCT 1320
DB 1261 TAGCCTATAGGTGTGGTATTACCACTATTAGCACTCCCTATTGAGTATAGCT 1320
QY 1321 TCCATTACTAATCCATAACATGCTCTTGGCCAACTATCTATTTGGCTATATGCCAA 1380
DB 1321 TCCATTACTAATCCATAACATGCTCTTGGCCAACTATCTATTTGGCTATATGCCAA 1380
QY 1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTAT 1440
DB 1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTAT 1440
QY 1441 TATTTACAAATTCACATATACAAACACGCGTCCCGGTGCCCGCAGTTTTTATTAACA 1500
DB 1441 TATTTACAAATTCACATATACAAACACGCGTCCCGGTGCCCGCAGTTTTTATTAACA 1500
QY 1501 TAGCGTGGGATCTCCGACATCTCGGTACGTGTTCGGACATGGGCTCTTCCCGTAGC 1560
DB 1501 TAGCGTGGGATCTCCGACATCTCGGTACGTGTTCGGACATGGGCTCTTCCCGTAGC 1560
QY 1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCTCGTCCAGCGGCTCATGGTCCGTCGGC 1620
DB 1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCTCGTCCAGCGGCTCATGGTCCGTCGGC 1620
QY 1621 AGCTCTCTGTCTCTAAACAGTGGAGCCAGACTTAGGCACAGCAATGCCACCAACC 1680
DB 1621 AGCTCTCTGTCTCTAAACAGTGGAGCCAGACTTAGGCACAGCAATGCCACCAACC 1680
QY 1681 AGTGTCCGCACAAAGCCGTGGCGGTAGGTATGTCTGAAATGAGCTCGAGATTGG 1740
DB 1681 AGTGTCCGCACAAAGCCGTGGCGGTAGGTATGTCTGAAATGAGCTCGAGATTGG 1740
QY 1741 GCTCGACCTGGACGAGATGGAAGACTTAAGGCAGCGGCAGAGAAGATGCGAGGAGCT 1800
DB 1741 GCTCGACCTGGACGAGATGGAAGACTTAAGGCAGCGGCAGAGAAGATGCGAGGAGCT 1800
QY 1801 GAGTGTGTGATTCTGATAAGAGTCAAGAGTAACTCCGTTGCGGTCTGTAAACGGTGG 1860
DB 1801 GAGTGTGTGATTCTGATAAGAGTCAAGAGTAACTCCGTTGCGGTCTGTAAACGGTGG 1860
QY 1861 AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACAGACATAATAGCT 1920
DB 1861 AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACAGACATAATAGCT 1920
QY 1921 GACAGACTAAACAGACTGTCTCTTCCATGGGTCTTTTCTGAGTCAACCGTCTGACGAA 1980
DB 1921 GACAGACTAAACAGACTGTCTCTTCCATGGGTCTTTTCTGAGTCAACCGTCTGACGAA 1980
QY 1981 TTCAAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAG 2040
DB 1981 TTCAAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAG 2040
QY 2041 CAGTCTTCGTTTGGCCAGCGCTAGCGAAACCCACGCTCACCGGGGGAAGTGCCTGCCACA 2100
DB 2041 CAGTCTTCGTTTGGCCAGCGCTAGCTA----- 2068
QY 2101 CTGTGTCTGGATTGTTAGCTCTCTCGCACAGCGGCCAAGCAGAACGCTCCAGCTGATCA 2160
DB 2069 -----CCAGTGGCCA 2079
QY 2161 ACACCAACCGGAGTGGCACTCAATAGACACGCGCCCTGAATGCAATGATAGCTCAACA 2220
DB 2080 ACAGCACCGGCTGTACCAAGTGA-----CCAAGCACTGCCCCCAACA 2121
QY 2221 CCGGCTGGTGGCAGGCGCTTTCTATCACCAAGTTCACTCTTACGAGCTGCTCCTGAGA 2280

DB 2122 GCAGCATCGTGTACAGAGCGCGCGACCATCTCTGCACACCCCGGCTGGTGCC----- 2176
QY 2281 GGCTAGCCAGCTGCGGACCCCTTACCGATTTTGAACAGGCTGGGGCCCTATACAGTTATG 2340
DB 2177 -----CTGCGTGGCGAGGGCAACGCGAGCGCTGCTGTGGTGGCCATGACCCCA 2226
QY 2341 CCAACGGAAGCGGCGCGACAGCGCCCTACTGTGGCACTACCCCCCAAAACCTTGGG 2400
DB 2227 CCGTGGCCACCCGCGACGGCAAGCTGCCCGCCACCCAGCTGCGCGC----- 2273
QY 2401 GTATTGTGCCCGGAAGTGTGTGTGTCGGTATATTGCTTCACTCCAGCCCGCTGG 2460
DB 2274 -----CACATCGACCTGC 2286
QY 2461 TGGTGGGAACGACGAGTCCGGCGCGCCACCTACAGTCTGGGTGAAATGATACGG 2520
DB 2287 TGGTGGGACGCGCACCCCTGTGCAGCGCC-----TGT 2319
QY 2521 AGTCTTCTGTCTTAAACAATACAGGCCACCGCTGGGCAATTTGGTTGGTTGTACTGGA 2580
DB 2320 AGTGGGCGACCTGTGCGCAGCGTGTCTGTGGGGCAGCTGTTC----- 2366
QY 2581 TGAACCTCAACTGGATTACCAAAAGTGTGGGAGCGCTCTCTGTGTCTATCGGAGGGCGG 2640
DB 2367 -----ACCT 2370
QY 2641 GCAACAACACCTGCATCGTCCCGCACATGATGTCTTCGGAAGCATCCGGAGCCACATCT 2700
DB 2371 TCAGCCCCCGCGCCACATGGACACCCAGGGCTGCAACTGCAGCATCTACCCCGGCCACA 2430
QY 2701 CTCGGTGGGCTCCGGTCCCTGGATCACACCCAGGTGCCTGGTTCGACTACCGGTATAGC 2760
DB 2431 TCACCGGCCACCGCATGGCTGGGACATGAT----- 2464
QY 2761 TTTGGCATTCTCTGTACCATCAACTACCATATTTAAATCAGGATGTACTGGGAG 2820
DB 2465 ----- 2464
QY 2821 GGGTCGAACACAGGCTGGAGTGCCTGCAACTGGACGGGGCGGAACTGTGGATCTGG 2880
DB 2465 ----- 2464
QY 2881 AAGATAGGACAGGTCGAGATCGATATGGAGAACATCACATCAGGATTCCTAGGACCC 2940
DB 2465 -----GAACTGGAGCCCGCACCAAGGAGAACATCACATCAGGATTCCTAGGACCC 2517
QY 2941 TGCTCGTGTACAGGCGGGTTTTCTTGTTCACAAGAAATCCTCAAAATACCGCAGAGTC 3000
DB 2518 TGCTCGTGTACAGGCGGGTTTTCTTGTTCACAAGAAATCCTCAAAATACCGCAGAGTC 2577
QY 3001 TAGACTCGGTGGAGTCTCTCAATTTCTAGGGGATCTCCCGTGTGTCTTGGCCAAA 3060
DB 2578 TAGACTCGGTGGAGTCTCTCAATTTCTAGGGGATCTCCCGTGTGTCTTGGCCAAA 2637
QY 3061 ATTGCGAGTCCCAACCTCCAATCACTCAACCTCTGCTCTCCAAATTTGCTGGT 3120
DB 2638 ATTGCGAGTCCCAACCTCCAATCACTCAACCTCTGCTCTCAATTTGCTGGT 2697
QY 3121 ATCGCTGGAGTGTCTGCGCGGTTTTATCATATTTCTTTCATCTCTGCTGCTATGCTCA 3180
DB 2698 ATCGCTGGAGTGTCTGCGCGGTTTTATCATATTTCTTTCATCTCTGCTGCTATGCTCA 2757
QY 3181 TCTTCTTATTGGTCTTCTGGATTATCAAGGTATGTTGCCGTTTGTCTCTTAATTCAG 3240
DB 2758 TCTTCTTATTGGTCTTCTGGATTATCAAGGTATGTTGCCGTTTGTCTCTTAATTCAG 2817
QY 3241 GATCAACAACCACTAGGGGACATGCAAAACCTGCAGCTCTCTGCTCAGGCAACT 3300
DB 2818 GATCAACAACAACCTAGGGGACATGCAAAACCTGCAGCTCTCTGCTCAGGCAACT 2877
QY 3301 CTATGTTTCCCTCATGTGCTGTACAAAACCTACGGATGGAATTCACCTGTATTCCCA 3360


```
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCAACCCAACTGATCTTC 5580
DB 5098 GATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCAACCCAACTGATCTTC 5157
QY 5581 AGCATCTTTTACTTTTACCATCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGC 5640
DB 5158 AGCATCTTTTACTTTTACCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGC 5217
QY 5641 AAAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTACTCTTCTCTTTTCAATA 5700
DB 5218 AAAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTACTCTTCTCTTTTCAATA 5277
QY 5701 TTATTGAAGCAATTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTA 5760
DB 5278 TTATTGAAGCAATTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTA 5337
QY 5761 GAAAAATAAACAATAAGGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGTCTA 5820
DB 5338 GAAAAATAAACAATAAGGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGTCTA 5397
QY 5821 AGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCCCTTTG 5880
DB 5398 AGAAACCATTTATTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCCCTTTG 5457
QY 5881 TC 5882
DB 5458 TC 5459

RESULT 3
AAD06791
ID AAD06791 standard; DNA; 5128 BP.
AC XX
AD AAD06791;
DT 06-AUG-2001 (first entry)
DE Plasmid pCMVII-ps2-SAg.
KW Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;
KW HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;
KW plasmid pCMVII-ps2-SAg; cyclic; circular; ds.
XX
OS Human cytomegalovirus.
OS Homo sapiens.
OS Hepatitis B virus.
OS Bos sp.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1988..2833
FT FT /*tag= a
FT FT /product= "pres2 and S domains of HBsAg"
FT FT misc_feature 1988..2152
FT FT /*tag= b
FT FT /note= "pres2 domain coding sequence"
FT FT misc_feature 2153..2833
FT FT /*tag= c
FT FT /note= "S domain coding sequence"
XX
PN WO200138358-A2.
XX
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032249.
XX
XX 24-NOV-1999; 99US-0167224P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Selby M, Glazer E, Houghton M;
XX
```

```
DR WPI; 2001-367661/38.
DR P-PSDB; AAS02620.
XX
PT Virus-like particle for use as an immunogen, comprising a first hepatitis
PT B virus surface antigen (HBsAg) and chimeric antigen comprising a second
PT HBsAg covalently linked to hepatitis C immunogenic polypeptide.
XX
PS Example 2; Fig 2; 115pp; English.
XX
CC The invention relates to a virus-like particle for use as an immunogen,
CC comprising a first hepatitis B virus surface antigen (HBsAg) and a
CC chimeric antigen comprising a second HBsAg which is covalently linked to
CC an hepatitis C virus (HCV) immunogenic polypeptide, where the first and
CC the second HBsAg each comprise a substantially complete S domain. The
CC virus-like particle is useful as immunogen and as vaccine. The present
CC sequence is plasmid pCMVII-ps2-SAg containing coding sequences for pres2
CC and S domains of HBsAg inserted into pCMVII between human cytomegalovirus
CC (CMV) intron A and bovine growth hormone poly A terminator (BGHT). The
CC plasmid pCMVII is a pUC19-based cloning vector designed for expression in
CC mammalian cells. It comprises human cytomegalovirus (CMV) immediate early
CC (IE) enhancer/promoter, human CMV intron A, a human tissue plasminogen
CC activator (tPA) leader, a bovine growth hormone poly A terminator (BGHT),
CC a ColE1 origin of replication and an ampicillin resistance (Amp R) gene.
CC The pCMVII-ps2-SAg plasmid is used for production of virus-like particle
CC of the invention
XX
SQ Sequence 5128 BP; 1218 A; 1321 C; 1249 G; 1340 T; 0 U; 0 Other;

Query Match 71.2%; Score 4190; DB 5; Length 5128;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 5038; Conservative 0; Mismatches 90; Indels 754; Gaps 4;

QY 1 TCGCGCGTTTCGGTGATGACGGGTGAAACCTCTGACACATGAGCTCCCGAGACGGTCA 60
DB 1 TCGCGCGTTTCGGTGATGACGGGTGAAACCTCTGACACATGAGCTCCCGAGACGGTCA 60
QY 61 CAGCTTGCTGTAAAGCGGATGCGGGGAGCAGACACAGCCGTCAGGCGCGTTCAGCGGGTG 120
DB 61 CAGCTTGCTGTAAAGCGGATGCGGGGAGCAGACACAGCCGTCAGGCGCGTTCAGCGGGTG 120
QY 121 TTGGCGGGGTGTCGGGGCTGGCTTAATCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB 121 TTGGCGGGGTGTCGGGGCTGGCTTAATCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
QY 181 ACCATATGAAGCTTTTTCGAAAAGCTAGGCTTCGAAAAGAGCTCTCACTACTTCTGG 240
DB 181 ACCATATGAAGCTTTTTCGAAAAGCTAGGCTTCGAAAAGAGCTCTCACTACTTCTGG 240
QY 241 AATAGCTCAGAGCGCGAGCGGCGCTCGGCTCTGCATTAATAAAAAAATTTAGTCAGCCA 300
DB 241 AATAGCTCAGAGCGCGAGCGGCGCTCGGCTCTGCATTAATAAAAAAATTTAGTCAGCCA 300
QY 301 TGGGCGGAGAAATGGCGGAACTGGGCGGGGAGGGAATTTATGGCTATTGGCCATTGCAAT 360
DB 301 TGGGCGGAGAAATGGCGGAACTGGGCGGGGAGGGAATTTATGGCTATTGGCCATTGCAAT 360
QY 361 AGTTGTATCTATATCATTAATATGATATTTATATTTGGCTCATGTCCCAATATGACGCCA 420
DB 361 AGTTGTATCTATATCATTAATATGATATTTATATTTGGCTCATGTCCCAATATGACGCCA 420
QY 421 TGTGACATTGATTATTGACTAGTTAATTAATAGTAATCAATTACGGGGTCATTAGTTTCAAT 480
DB 421 TGTGACATTGATTATTGACTAGTTAATTAATAGTAATCAATTACGGGGTCATTAGTTTCAAT 480
QY 481 AGCCCATATATGGAGTTCCGCGTTACATAAATTTACGTAATAATGGCCCGCTGGCTGACCG 540
DB 481 AGCCCATATATGGAGTTCCGCGTTACATAAATTTACGTAATAATGGCCCGCTGGCTGACCG 540
QY 541 CCCAACGACCCCGCCCATTTGACGTCATTAATGACGATATGTTCCCAATAGTAAGCCCAATA 600
DB 541 CCCAACGACCCCGCCCATTTGACGTCATTAATGACGATATGTTCCCAATAGTAAGCCCAATA 600
QY 601 GGGACTTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 660
```


QY	2821	GGTCCGAACACAGCGCTGGAAAGCTGCCCTGCACACTGGACGCGGGGCGGAACGTTTGGATCTGTG	2880
Db	2142	-----	2141
QY	2881	AAGATAGGACAGGTCGAGATCGATATGAGAACATCATCAGGATTCCTAGGACCCC	2940
Db	2142	-----	2141
QY	2941	TGCTCGTGTACAGCGGGGTTTTCTTGTGTGACAGAAATCCCTCACAAATACCCGAGAGTC	3000
Db	2187	TGCTCGTGTACAGCGGGGTTTTCTTGTGTGACAGAAATCCCTCACAAATACCCGAGAGTC	2246
QY	3001	TAGACTCGTGGTGGACTTCTCTCAATTTCTAGGGGATCTCCGCTGTGTCTTGGCCAAA	3060
Db	2247	TAGACTCGTGGTGGACTTCTCTCAATTTCTAGGGGATCTCCGCTGTGTCTTGGCCAAA	2306
QY	3061	ATTTCGAGTCCCAACCTCCAATCACTCACCAACCTCCCTGCTCCCAATTTGTCTGGTT	3120
Db	2307	ATTTCGAGTCCCAACCTCCAATCACTCACCAACCTCCCTGCTCCCAATTTGTCTGGTT	2366
QY	3121	ATCGCTGGATGTCTCGCGGTTTTATCATATTTCTCTTCATCTCTGCTGCTATGCTCA	3180
Db	2367	ATCGCTGGATGTCTCGCGGTTTTATCATATTTCTCTTCATCTCTGCTGCTATGCTCA	2426
QY	3181	TCCTCTTATTTGGTTCTTCTGGATATCAAGGTATGTTGCCGTTGTCTCTAATTTCCAG	3240
Db	2427	TCCTCTTATTTGGTTCTTCTGGATATCAAGGTATGTTGCCGTTGTCTCTAATTTCCAG	2486
QY	3241	GATCAACAACAAACAGTACGGGACCATGCAAAACCTGCAGACTCCCTGCTCAAGGCAACT	3300
Db	2487	GATCAACAACAAACAGTACGGGACCATGCAAAACCTGCAGACTCCCTGCTCAAGGCAACT	2546
QY	3301	CTATGTTTCCCTCATGTTGCTGACAAAACCTACGGATGGAATATGCACCTGTATTTCCCA	3360
Db	2547	CTATGTTTCCCTCATGTTGCTGACAAAACCTACGGATGGAATATGCACCTGTATTTCCCA	2606
QY	3361	TCCCATGCTCGGCTTTTCGAAAATACCTATGGAGTGGGCTCAGTCCGTTCTCTT	3420
Db	2607	TCCCATGCTCGGCTTTTCGAAAATACCTATGGAGTGGGCTCAGTCCGTTCTCTT	2666
QY	3421	GGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTTCGTAGGGCTTTCCCCACHTTTTGGC	3480
Db	2667	GGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTTCGTAGGGCTTTCCCCACHTTTTGGC	2726
QY	3481	TTTCAGCTATATGGATGATGTGTTATTTGGGGGCCAAGTCTGTACAGCATCGTAGTCCCT	3540
Db	2727	TTTCAGCTATATGGATGATGTGTTATTTGGGGGCCAAGTCTGTACAGCATCGTAGTCCCT	2786
QY	3541	TTATACCGCTGTTACCAATTTCTTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG	3600
Db	2787	TTATACCGCTGTTACCAATTTCTTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG	2846
QY	3601	AGCAAGTCTAGAAAGCGCCCAAGATATCAAGGATCCACTACGCTTAGAGCTCGCTGA	3660
Db	2847	AGCAAGTCTAGAAAGCGCCCAAGATATCAAGGATCCACTACGCTTAGAGCTCGCTGA	2906
QY	3661	TCAGCCTCGACTGTGCCCTTCTAGTTGCGAGCCATCTGTTGTTTGGCCCTCCCGCTGCCT	3720
Db	2907	TCAGCCTCGACTGTGCCCTTCTAGTTGCGAGCCATCTGTTGTTTGGCCCTCCCGCTGCCT	2966
QY	3721	TCCTTGACCTGGAAAGTGCCACTCCCACTGTCTTTCTTCTTAATAAATGAGGAAATTGCA	3780
Db	2967	TCCTTGACCTGGAAAGTGCCACTCCCACTGTCTTTCTTCTTAATAAATGAGGAAATTGCA	3026
QY	3781	TCCATTTGCTGAGTAGTGTCTATTTCTGGGGGTTGGGTTGGGAGGACAGCAAG	3840
Db	3027	TCCATTTGCTGAGTAGTGTCTATTTCTGGGGGTTGGGTTGGGAGGACAGCAAG	3086
QY	3841	GGGAGGATTTGGGAAGCAATAGCAGGACATGCTGGGAGCTCTTCCGCTTCTCGCTCAC	3900
Db	3087	GGGAGGATTTGGGAAGCAATAGCAGGACATGCTGGGAGCTCTTCCGCTTCTCGCTCAC	3146

QY	3901	TGACTCGCTCGCTCGGTCGTTCCGCTCGCGAGCGGTATCAGTCACTCAAGCGGT	3960
Db	3147	TGACTCGCTCGCTCGGTCGTTCCGCTCGCGAGCGGTATCAGTCACTCAAGCGGT	3206
QY	3961	AATACCGTTTATCCACAGAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCA	4020
Db	3207	AATACCGTTTATCCACAGAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCA	3266
QY	4021	GCAAAAGGCCAGAACCCGTAAAGCCGCTGCTGGCGTTTTCATAGGCTCCGCC	4080
Db	3267	GCAAAAGGCCAGAACCCGTAAAGCCGCTGCTGGCGTTTTCATAGGCTCCGCC	3326
QY	4081	CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGACT	4140
Db	3327	CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGACT	3386
QY	4141	ATAAAGATACCAAGCGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTTCCGACCT	4200
Db	3387	ATAAAGATACCAAGCGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTTCCGACCT	3446
QY	4201	GCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATG	4260
Db	3447	GCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATG	3506
QY	4261	CTCAGCTCTAGGTATCTCAGTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCA	4320
Db	3507	CTCAGCTCTAGGTATCTCAGTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCA	3566
QY	4321	CGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAGTCCAA	4380
Db	3567	CGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAGTCCAA	3626
QY	4381	CCCGTAAACACACGACTTATTCGCCACTGGGACGAGCACCTGGTAAACAGGATAGCAGAGC	4440
Db	3627	CCCGTAAACACACGACTTATTCGCCACTGGGACGAGCACCTGGTAAACAGGATAGCAGAGC	3686
QY	4441	GAGTATGTAGGGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAG	4500
Db	3687	GAGTATGTAGGGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAG	3746
QY	4501	AAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGAGTTGG	4560
Db	3747	AAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGAGTTGG	3806
QY	4561	TAGCTCTTGAATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGTCAGCA	4620
Db	3807	TAGCTCTTGAATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGTCAGCA	3866
QY	4621	GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTTCTACGGGTC	4680
Db	3867	GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTATCTTTTCTACGGGTC	3926
QY	4681	TGACGCTCAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAATGAGATTATCAAAAG	4740
Db	3927	TGACGCTCAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAATGAGATTATCAAAAG	3986
QY	4741	GATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCTAAGATATATA	4800
Db	3987	GATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCTAAGATATATA	4046
QY	4801	TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGGACCTTATCTCAGGAT	4860
Db	4047	TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGGACCTTATCTCAGGAT	4106
QY	4861	CTGCTCTATTTCTGTTTCAATAGTTCCTGACTCCCGCTCGTGTAGATAACTACGATACG	4920
Db	4107	CTGCTCTATTTCTGTTTCAATAGTTCCTGACTCCCGCTCGTGTAGATAACTACGATACG	4166
QY	4921	GGAGGCTTACCAATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACGGC	4980
Db	4167	GGAGGCTTACCAATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACGGC	4226
QY	4981	TCCAGATTTTACAGCAATAAACCCAGCCAGCGGAGCGGAGAGTGTGCTCTGC	5040

```
Db 4227 TCAGATTTATCAGCAATAAACAGCAGCAGCGGAAGGGCGAGCAGAAAGTGGTCTGC 4286
Qy 5041 AACTTTATCCGCTCCATCAGTCTATTAACTTCCCGGAGCTAGAGTAGTAGTTC 5100
Db 4287 AACTTTATCCGCTCCATCAGTCTATTAACTTCCCGGAGCTAGAGTAGTAGTTC 4346
Qy 5101 GCCAGTTAATAGTTTCCGCAACGTTTGTGCAATGTCTACAGGCATCGTGTGTACGGCTC 5160
Db 4347 GCCAGTTAATAGTTTCCGCAACGTTTGTGCAATGTCTACAGGCATCGTGTGTACGGCTC 4406
Qy 5161 GTCGTTTGTATGGCTTCATTCAGTCCCGTTCCTCAACGATCAAGCGAGTTACATGATC 5220
Db 4407 GTCGTTTGTATGGCTTCATTCAGTCCCGTTCCTCAACGATCAAGCGAGTTACATGATC 4466
Qy 5221 CCCCATGTTGTGCAAAAGCGGTAGTCTCTTCGGTCCCTCCGATCGTGTGTACAGAGTAA 5280
Db 4467 CCCCATGTTGTGCAAAAGCGGTAGTCTCTTCGGTCCCTCCGATCGTGTGTACAGAGTAA 4526
Qy 5281 GTTGGCGCAGTGTATCACTCATGTGTTATGCGCAGCACTGATATAATTCCTTACTGTCAAT 5340
Db 4527 GTTGGCGCAGTGTATCACTCATGTGTTATGCGCAGCACTGATATAATTCCTTACTGTCAAT 4586
Qy 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATTCGTAGAATA 5400
Db 4587 GCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATTCGTAGAATA 4646
Qy 5401 GTGTATGCGCGCAGCGAGTGTCTTTCGCGCGGTCAATACGGGTAATACCGGCCACA 5460
Db 4647 GTGTATGCGCGCAGCGAGTGTCTTTCGCGCGGTCAATACGGGTAATACCGGCCACA 4706
Qy 5461 TAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTCTTCGGGCGGAAACCTCTCAAG 5520
Db 4707 TAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTCTTCGGGCGGAAACCTCTCAAG 4766
Qy 5521 GATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCAACTGATCTTC 5580
Db 4767 GATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCAACTGATCTTC 4826
Qy 5581 AGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATCCGC 5640
Db 4827 AGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATCCGC 4886
Qy 5641 AAAAAAGGGAATAAGCGGCACACGGAAATGTTCAATACTCATACTCTTCTTTTCAATA 5700
Db 4887 AAAAAAGGGAATAAGCGGCACACGGAAATGTTCAATACTCATACTCTTCTTTTCAATA 4946
Qy 5701 TTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
Db 4947 TTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTA 5006
Qy 5761 GAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAGATGCCACCTGACGCTTA 5820
Db 5007 GAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGAAGATGCCACCTGACGCTTA 5066
Qy 5821 AGAAACCATTTATTCATGACATTAACCTATAAAATAGCGGTATCAGGAGGCCCTTTTCG 5880
Db 5067 AGAAACCATTTATTCATGACATTAACCTATAAAATAGCGGTATCAGGAGGCCCTTTTCG 5126
Qy 5881 TC 5882
Db 5127 TC 5128
```

RESULT 4

AAF83669

ID AAF83669 standard; DNA; 6299 BP.

XX

AC AAF83669;

XX

DT 23-JUL-2001 (first entry)

XX

DE HCV NS34A ORF comprising pCMV-NS34A nucleic acid sequence.

```
XX HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW immunotherapy; NS34A; ds.
XX Synthetic.
OS Hepatitis C virus.
XX Key Location/Qualifiers
FH 1990..4050
FT /*tag= a
FT /product= "NS34A"
XX
XX WO200138360-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032326.
XX
XX 24-NOV-1999; 99US-0167502P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Coit D, Medina-Selby A, Selby M, Houghton M;
PI WPI; 2001-343948/36.
XX
XX P-PSDB; AAB62633.
XX
XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT a vaccine against HCV, comprises a polypeptide having a mutation that
PT functionally disrupts the catalytic domain of NS3.
XX
XX Disclosure; Fig 9; 340pp; English.
XX
XX The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC the catalytic domain of NS3, where the mutation functionally disrupts the
CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC in vaccines against HCV and as diagnostics. The antibodies raised against
CC these polypeptides can also be used as diagnostics, or for passive
CC immunotherapy. The antibodies are also useful for isolating and
CC identifying HCV particles. The present sequence represents the nucleic
CC acid sequence of the pCMV-NS34A comprising the ORF encoding a NS34A
CC polypeptide
XX
SQ Sequence 6299 BP; 1449 A; 1713 C; 1637 G; 1500 T; 0 U; 0 Other;
```

Query Match 63.1%; Score 3713.6; DB 4; Length 6299;

Best Local Similarity 77.4%; Pred. No. 0;

Matches 4876; Conservative 0; Mismatches 1004; Indels 420; Gaps 14;

Qy 2 CGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGACGCTCCCGAGACGGTCA 61

Db 1 CGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGACGCTCCCGAGACGGTCA 60

Qy 62 AGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCCGTCAGGCGCGCTCAGCGGGTGT 121

Db 61 AGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCCGTCAGGCGCGCTCAGCGGGTGT 120

Qy 122 TGGCGGGTGTGCGGGCTGGCTTAACTATTCGGCATCAGAGCAGATGTGTAAGAGTGCA 181

Db 121 TGGCGGGTGTGCGGGCTGGCTTAACTATTCGGCATCAGAGCAGATGTGTAAGAGTGCA 180

Qy 182 CCATATGAAGCTTTTGGAAAAAGCGCTAGGCCTCCAAAAAGCGCTCTCTACTTCTGGA 241

Db 181 CCATATGAAGCTTTTGGAAAAAGCGCTAGGCCTCCAAAAAGCGCTCTCTACTTCTGGA 240

Qy 242 ATAGCTCAGAGCGCGGCGGCTCGGCTCTGTCATAAATAAAAAAATAGTCAGCCAT 301

Db 241 ATAGCTCAGAGCGCGGCGGCTCGGCTCTGTCATAAATAAAAAAATAGTCAGCCAT 300

QY 302 GGGCGGAGAAATGGCGGAACCTGGCGGGAGGAAATTATTGGCTATTGGCCATTGCATA 361
DB |||||
QY 301 GGGCGGAGAAATGGCGGAACCTGGCGGGAGGAAATTATTGGCTATTGGCCATTGCATA 360
DB |||||
QY 362 CGTTGATCTATATCATAAATATGTACATTTTATATGGCTCATGTCCAATATGACCGCCAT 421
DB |||||
QY 361 CGTTGATCTATATCATAAATATGTACATTTTATATGGCTCATGTCCAATATGACCGCCAT 420
DB |||||
QY 422 GTTGACATTTGATTTAGCTAGTTATTAATAGTAATAAATTAACGAGGTCATTTAGTTCA 481
DB |||||
QY 421 GTTGACATTTGATTTAGCTAGTTATTAATAGTAATAAATTAACGAGGTCATTTAGTTCA 480
DB |||||
QY 482 GCCCATATATGAGTTCCGGTTACATAAATTAACGTTAAATGACCGCTGGCTGACCGC 541
DB |||||
QY 481 GCCCATATATGAGTTCCGGTTACATAAATTAACGTTAAATGACCGCTGGCTGACCGC 540
DB |||||
QY 542 CCAACGACCCCGCCCATTAACGCTCAATAATAGCATGTATGTTCCCATAGTAACCGCCCAATAG 601
DB |||||
QY 541 CCAACGACCCCGCCCATTAACGCTCAATAATAGCATGTATGTTCCCATAGTAACCGCCCAATAG 600
DB |||||
QY 602 GGACTTTCCATTTGACGTCGAATGGGTGGAGTATTTACGTTAAACTGCCCACTTGGCAGTAC 661
DB |||||
QY 601 GGACTTTCCATTTGACGTCGAATGGGTGGAGTATTTACGTTAAACTGCCCACTTGGCAGTAC 660
DB |||||
QY 662 ATCAAGTATCATATATGCCAAGTCCGCGCCCTATTGACGTCATATGACGTTAAATGACCGC 721
DB |||||
QY 661 ATCAAGTATCATATATGCCAAGTCCGCGCCCTATTGACGTCATATGACGTTAAATGACCGC 720
DB |||||
QY 722 CTTGGCATTTATGCCAGTACATGACCTTACGGACTTTTCTACTTGGCAGTACATCTACG 781
DB |||||
QY 721 CTTGGCATTTATGCCAGTACATGACCTTACGGACTTTTCTACTTGGCAGTACATCTACG 780
DB |||||
QY 782 TATTAGTCATCGCTATTACATCGGTGATGCGGTTTTTGGCAGTACACCAATGGCGGTGGAT 841
DB |||||
QY 781 TATTAGTCATCGCTATTACATCGGTGATGCGGTTTTTGGCAGTACACCAATGGCGGTGGAT 840
DB |||||
QY 842 AGCGGTTTGACTCAGCGGGATTTCCAGTCTCCACCCCATTTGACGTCMAATGGGAGTTGT 901
DB |||||
QY 841 AGCGGTTTGACTCAGCGGGATTTCCAGTCTCCACCCCATTTGACGTCMAATGGGAGTTGT 900
DB |||||
QY 902 TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTGTAATAACCGCCCGCCCTTGCACGC 961
DB |||||
QY 901 TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTGTAATAACCGCCCGCTTGCACGC 960
DB |||||
QY 962 AAATGGCGGTAGCGGTGATACGTTGGAGGTCTATATAGCAGAGCTCGTTTATGTAACC 1021
DB |||||
QY 961 AAATGGCGGTAGCGGTGATACGTTGGAGGTCTATATAGCAGAGCTCGTTTATGTAACC 1020
DB |||||
QY 1022 GTCAGATCGCTCGAGACGCCATCCACGCTGTTTTCACCTCCATAGAAGACACCGGGACC 1081
DB |||||
QY 1021 GTCAGATCGCTCGAGACGCCATCCACGCTGTTTTCACCTCCATAGAAGACACCGGGACC 1080
DB |||||
QY 1082 GATCCAGCCTTCGCGCGCGGGAACGGTGCAATTGGAACGGGGAATCCCGCTGCCAAGAGTG 1141
DB |||||
QY 1081 GATCCAGCCTTCGCGCGCGGGAACGGTGCAATTGGAACGGGGAATCCCGCTGCCAAGAGTG 1140
DB |||||
QY 1142 ACCTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATAC 1201
DB |||||
QY 1141 ACCTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATAC 1200
DB |||||
QY 1202 TGTTTTGGCTTGGGCGCTATACACCCCGCTCTTATGCTATAGGTGATGTTAGCTT 1261
DB |||||
QY 1201 TGTTTTGGCTTGGGCGCTATACACCCCGCTCTTATGCTATAGGTGATGTTAGCTT 1260
DB |||||
QY 1262 AGCCTATAGGTGTGGGTTATTGACCAATTTATGACCACTCCCTCTATTGGTGACCATCTTT 1321
DB |||||
QY 1261 AGCCTATAGGTGTGGGTTATTGACCAATTTATGACCACTCCCTCTATTGGTGACCATCTTT 1320
DB |||||
QY 1322 CCATTAATCAATACATGGCTCTTTGGCACAACATCTCTATTGGCTATATGCCAAT 1381
DB |||||
QY 1321 CCATTAATCAATACATGGCTCTTTGGCACAACATCTCTATTGGCTATATGCCAAT 1380
DB |||||
QY 1382 ACTCTGTCTTACAGACTTGACAGGACTCTGTATTTTACAGGATGGGGTCCATTATT 1441
DB |||||

DB |||||
QY 1381 ACTCTGTCTTACAGACTTGACACGAGCTCTGTATTTTACAGATGGGTCATTATT 1440
DB |||||
QY 1442 ATTTACAAATTCACATATACAAACCGCGTCCCGCTGCCCGCAGTTTTTATTAACAT 1501
DB |||||
QY 1441 ATTTACAAATTCACATATACAAACCGCGTCCCGCTGCCCGCAGTTTTTATTAACAT 1500
DB |||||
QY 1502 AGGTTGGGATTCCTCCGACATCTCGGTACCGTTCGGGACATCGGGCTCTTCTCCGGTAGCG 1561
DB |||||
QY 1501 AGGTTGGGATTCCTCCGACATCTCGGTACCGTTCGGGACATCGGGCTCTTCTCCGGTAGCG 1560
DB |||||
QY 1562 GGGAGCTTCCACATCCGAGCCCTGCTCCATCCGTCAGCGCTCATGCTCGCTCGGCA 1621
DB |||||
QY 1561 GGGAGCTTCCACATCCGAGCCCTGCTCCATCCGTCAGCGCTCATGCTCGCTCGGCA 1620
DB |||||
QY 1622 GCTCTTGTCTCTTAACAGTGGAGGCCAGACTTAGGCACAGCACAATGCCACCAACCA 1681
DB |||||
QY 1621 GCTCTTGTCTCTTAACAGTGGAGGCCAGACTTAGGCACAGCACAATGCCACCAACCA 1680
DB |||||
QY 1682 GTGTGCCGACAAAGCCGTGGGTAGGATATGTCTGAAAATAGCTCGGAGATTGG 1741
DB |||||
QY 1681 GTGTGCCGACAAAGCCGTGGGTAGGATATGTCTGAAAATAGCTCGGAGATTGG 1740
DB |||||
QY 1742 CTCGCACCTGGACGAGATGGAAGACTTAAGGCAGCGGCAGAGAAGATGTCAGGAGCTG 1801
DB |||||
QY 1741 CTCGCACCTGGACGAGATGGAAGACTTAAGGCAGCGGCAGAGAAGATGTCAGGAGCTG 1800
DB |||||
QY 1802 AGTTGTTGATTTCTGATAAGTCAAGATACTCCCGTTGCGGTCTGTTAACCGTGA 1861
DB |||||
QY 1801 AGTTGTTGATTTCTGATAAGTCAAGATACTCCCGTTGCGGTCTGTTAACCGTGA 1860
DB |||||
QY 1862 GGGCAGTGTAGTCTGAGCAGTACTCGTCTGCGCGCGCGCCACAGACATAATAGCTG 1921
DB |||||
QY 1861 GGGCAGTGTAGTCTGAGCAGTACTCGTCTGCGCGCGCGCCACAGACATAATAGCTG 1920
DB |||||
QY 1922 ACAGACTAACAGACTGTTCTTTTCCATGGGTCTTTTCTGAGTCAACCGCTGTCAGCAAT 1981
DB |||||
QY 1921 ACAGACTAACAGACTGTTCTTTTCCATGGGTCTTTTCTGAGTCAACCGCTGTCAGCTAA 1980
DB |||||
QY 1982 TCAAGCAATCATGGATGCAATGAAGAGGGCTCTGCTGTGTGCTG----- 2027
DB |||||
QY 1981 GAATTCACCATGGCGCCATCAGCGGTACGCGCCAGCAGACAAGGGGCTCTCCTAGGCTG 2040
DB |||||
QY 2028 -----CTGCTGTGTGGACAGTCTTCGTTTCG 2054
DB |||||
QY 2041 ATAATCACCAGCCTAATGGCGGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTG 2100
DB |||||
QY 2055 CCAGCGCTTAGCGAAAC-----C 2072
DB |||||
QY 2101 TCAACTGCTGCCCAACCTTCTTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTAC 2160
DB |||||
QY 2073 CACGTCAACCGGGGAGTCCGGCCACACTGTGTCTGGATTTGTTAGCTCTCCTCGCACCA 2132
DB |||||
QY 2161 CACGGGGCGGAACGAGGACCATCGGTCAACCAAGGTCTGTCTCATCAGATGTATACC 2220
DB |||||
QY 2133 GGGGCCAACAGAACGT-----CCAGCTGATCAACACCA 2166
DB |||||
QY 2221 AATGTAGACCAAGACCTTGTGGGCTGGCCCGCTTCGCAAGGTACCCGCTCATTTGACACC 2280
DB |||||
QY 2167 AC-----GGCAGTTGGCACCTCAATAGCACGGCCCTG 2198
DB |||||
QY 2281 TGCATTTGGGCTCTCGGACCTTTACTGTGTCAAGAGGACGCCGATGTCTATTCGGTG 2340
DB |||||
QY 2199 AACTGCAATGATAGCCTCAACACCGGCTGGTTGGAGGGCTTTTATACCAACAAGTTC 2258
DB |||||
QY 2341 CGCGCGCGGGGTGATAGCAGGGGACAGCTGTCTGCGCCCGGCCATTTCTACTTTGAAA 2400
DB |||||
QY 2259 AACTCTTCAAGGCTGCTGAGAGGCTAGCACCTGCCGACCCCTTACCCGATTTTGACAG 2318
DB |||||
QY 2401 GGTCTCTCGGGGGGTCCGCTGTG-TGCCCGCGGGGACAGCGGTGGGCATATTTAGGGC 2459
DB |||||
QY 2319 GGTGGGGGCCCTATCAGTTTATGCCAAGGAAAGCGGCCCGGACAGCGCCCTACTGCTGG 2378
DB |||||

Db 2460 CGCGGTGTGCACCGTGGAGTGGCTAAGCGGTGGACTTTTATCCCTGTGGAGAACCTAGA 2519
QY 2379 CACTACCCCCAAAACCTTGGCGT-----ATTGTGCCCGGAAGAGTGTGTGGT 2429
Db 2520 GACAAACATGAGTCCCGGTGTACGGATACTCTCTCCACCACTAGTAGTCCCGCAGAG 2579
QY 2430 CGGGTATATGTCTTCACTCCAGCCCGGTGGTGGGAACGACCAAGGTCCGGCGCG 2489
Db 2580 CTTCCAGGTGGCTCACTCCATGCTCCACAGGCGGGCAAGACCAAGGTCCCGGC 2639
QY 2490 CCACCTACAGCTGGG-----TGAAAATGATACGGA 2521
Db 2640 TGCAATGACAGCTCAGGGCTAAGGTGTAGTACTCAACCCCTCTGTGTCTCAACACT 2699
QY 2522 CGTCTTGTCTTAAACAATACCAAGGCACCGCTGGGCAATGGTTCGGTGTACCTGGAT 2581
Db 2700 GGGCTTGTGTCTTACATGTCGAAGGTATGGGATCGATCTAAATCATCAGGACCGGGT 2759
QY 2582 GAATCAACT----- 2591
Db 2760 GAGAAACAAATTACCACTGGCAGCCCCATCACGTACTCCACCTACGGCAAGTTCTTCCCGA 2819
QY 2592 -----GGATTCACAAAGTGTGGAGGGC----- 2617
Db 2820 CGCGGGTGTCTCGGGGGCGCTTATGACATAATATTTGTGACGAGTGCCACTCCACGGA 2879
QY 2618 TCCTGTGTCTATCGAGGGCGGGCAACAAACCCCTGCACCTGCCCACTGATTGCTTCG 2677
Db 2880 TGCACATCATTTGGGCATTTGGCACTGTCTTTGACCAAGAGACTGCGGGGGGAG 2939
QY 2678 CAAGCATCCGAGCCACATPACTCTGGTGGCTCCGGTCCCTGGATCACACCCAGGTG 2737
Db 2940 ACTGGTGTGCTGCCACGCCACCCCTCGGGCTCGTCACTGTGCCCATCCCAACAT 2999
QY 2738 CTGTGTGACTACCGGTATAGGCTTTGGCATATCTTTGTACCATCAACTACACATATT 2797
Db 3000 CGAGGAGTGTCTGTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCT 3059
QY 2798 TAAAAATCAGGATGTAGTGGGGGGTTCGAACACAGGC--TGAAGCTGCTGCAACTGG 2855
Db 3060 CGAAGTAATCAGGGGGGAGACATCTCATCTTCTGTCTTCAAGAAAGAGTGGACGA 3119
QY 2856 ACGCGGGGCGAA--CGTTGCGATCTGGAAGATAGGACAGGTCCGAGATGATATGGAGA 2913
Db 3120 ACTCGCGCAAAAGCTGTGCAATTGGGCATCAATGCGGTGCTACTACCGGGTCTTGA 3179
QY 2914 ACATCATCAGGATTCCTAGGACCCCTGCTGTGTACAGGCGGGTCTTCTGTGGA 2973
Db 3180 CGTGTCCGTCACTCCGACACGCGCGATGTTGTGTCGTGGCAACCGATGCCCTCATGAC 3239
QY 2974 CAAGAATCCTCAATACCGCAGAG-----TCTAGACTGCTGTGGACTTCTCTC 3023
Db 3240 CGGCTATACCGGCGACTTCCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGT 3299
QY 3024 AATTTTCTTAGGGGATCTCCCGTGTGTCTTGGCCAAAATTCGAGTCCCCCAACCTCCAAT 3083
Db 3300 CGATTTACGCTTGACCTTACCATTTGACAAATCAGCAATCACGCTCCCCCAAGATGCTGT 3359
QY 3084 CATCAACAACTCTGTCTCCAAATTTGCTGTGTATCGCTGGATGTGTCTGGCGGT 3143
Db 3360 CTCCCGCACTCAACGTCGGGGGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTGT 3419
QY 3144 TTTATCATATCTCTTCACTCTGCTATGCTCATCTTCTTATTGTTCTTCTGGAT 3203
Db 3420 GGCACCGGGGGAGGCCCTCCGGCATGTTCGACTGTGCTCGTCTCTGTGAGTGTATGA 3479
QY 3204 TATCAAGGTATGTTGGCCGTTTGTCTCTAATTTCCAGGATCAACAACACAGTACGGGA 3263
Db 3480 CGCAGGCTGTGTTGATAGCTCAGCCCGCGGAGACTACAGTTAGGCTACAGCGTA 3539
QY 3264 CCATGCAAAACCTGACAGACTCTCTGTCTCAAGGCAACTCT----- 3302
Db 3540 CATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTT 3599

QY 3303 ----ATGTTTCCCTCATGTGTGTGTAACAAACCTACGATGGAATTCACCTGTATTCC 3359
Db 3600 TACAGGCTCACTCATATAGATGCCACCTTTCTATCCAGACAAACAGAGTGGGAGAA 3659
QY 3360 ATCCCATCGTCTCGGCTTTCGCAAAATACCTATGGAGTGGGCTCAGTTCGTTCTCT 3419
Db 3660 CTTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGTAGGGCTCAAGCCCCCTCCC 3719
QY 3420 TGGCTCAGTTTACTAGTGCCATTTGTTTCAAGTGTTCGTAGGGCTTCCCCCACTGTTGG 3479
Db 3720 ATCGTGGGACCAAGATGTGGAAGTGTGTTGCTTGCCTCAAGCCACCTCATGGGCAAC 3779
QY 3480 CTTTTCAGCTATATGGAATGTGGAATTTGGGGGCCAAAGTCTGTACAGCATCTGTAGTCCC 3539
Db 3780 ACCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAATCAACCTGACGACCCAGTCA 3639
QY 3540 TTTATAC----- 3547
Db 3840 CAAATACATCATGACATGATGTCGCGCCGACCTGGAGGTCTGCACGAGCACCTGGGTGCT 3899
QY 3548 ----- 3547
Db 3900 CGTGGCGGCTCTGGCTGCTTTTGGCCGCGTATTGCTGTCAACAGGCTGCGTGTGTCAT 3959
QY 3548 ----GCTGTACCAATTTTCTTCTCTGGGTATACATTTAAGAAATTCAGACTCGAG 3602
Db 3960 AGTGGGCGGCTGCTTCTGTCGGGAAGCCGCAATCATACCTGACAGGAAGTCTCTTA 4019
QY 3603 CAAAGTCTAAGAGGCGCGCAAGATATCAAGATTCACCTACGGGTGTAGAGTCTCGTATC 3662
Db 4020 CCGAGAGTTCGATAGATGGAAGAGTGTAGGATCCACTACGCGTGTAGAGTCTCGTATC 4079
QY 3663 AGCTCGACTGTCCTTCTAGTTGCCAGCATCTGTTGTCCTCCCTCCCGTGGCTTC 3722
Db 4080 AGCTCGACTGTCCTTCTAGTTGCCAGCATCTGTTGTTGCCCTTCCCGTGGCTTC 4139
QY 3723 CTTGACCTTGAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGAAATTCATC 3782
Db 4140 CTTGACCTTGAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGAAATTCATC 4199
QY 3783 GATTTGTGTAGTAGTGTCTATTCTTGGGGGTGGGGTGGGCGAGGACAGCAAGG 3842
Db 4200 GATTTGTGTAGTAGTGTCTATTCTTGGGGGTGGGGTGGGCGAGGACAGCAAGG 4259
QY 3843 GAGGATTCGGGAAGCAATAGCAGCATCTCGGGAGCTTTCGCTCTCTCGCTCACTG 3902
Db 4260 GGAGGATTCGGGAAGCAATAGCAGCATCTCGGGAGCTTTCGCTCTCTCGCTCACTG 4319
QY 3903 ACTCGCTGCTCGGTCTGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAA 3962
Db 4320 ACTCGCTGCTCGGTCTGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAA 4379
QY 3963 TACGTTATTCACAGAAATCAGGGAATAACGAGGAAGAAACATGTAGCAAAAGGCGAGC 4022
Db 4380 TACGTTATTCACAGAAATCAGGGAATAACGAGGAAGAAACATGTAGCAAAAGGCGAGC 4439
QY 4023 AAAAGGCGAGGAACCGTAAAGGCGCGTTCGCTGGCGTTTTCATAGGCTCCGCCCC 4082
Db 4440 AAAAGGCGAGGAACCGTAAAGGCGCGTTCGCTGGCGTTTTCATAGGCTCCGCCCC 4499
QY 4083 CTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGGGGAACCCGACAGGACTAT 4142
Db 4500 CTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGGTGGGGAACCCGACAGGACTAT 4559
QY 4143 AAAGATAACAGGCTTTCCTTGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCTGTC 4202
Db 4560 AAAGATAACAGGCTTTCCTTGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCTGTC 4619
QY 4203 CGCTTACCGATACCTGTCGCTTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCT 4262
Db 4620 CGCTTACCGATACCTGTCGCTTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCT 4679

QY 4263 CACGCTGATAGTATCTCAGTTCCGTTAGGTAGTTCGCTCCAGCTGGGCTGTGTCACG 4322
DB |||||
QY 4680 CACGCTGATAGTATCTCAGTTCCGTTAGGTAGTTCGCTCCAGCTGGGCTGTGTCACG 4739
DB |||||
QY 4323 AACCCCGGTTACGCGCCGACCGCTGGCCCTTATCCGGTAACCTATCGCTTTAGTCCAAAC 4382
DB |||||
QY 4740 AACCCCGGTTACGCGCCGACCGCTGGCCCTTATCCGGTAACCTATCGCTTTAGTCCAAAC 4799
DB |||||
QY 4383 CGGTAAAGACAGCTTATCCCACTACGCGACAGCCACTGGTAAACAGATTTAGCAGAGCA 4442
DB |||||
QY 4800 CGGTAAAGACAGCTTATCCCACTACGCGACAGCCACTGGTAAACAGATTTAGCAGAGCA 4859
DB |||||
QY 4443 GGTATCTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGCTACACTAGAA 4502
DB |||||
QY 4860 GGTATCTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTAACTACGCTACACTAGAA 4919
DB |||||
QY 4503 GGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGTAACTTCGGAAAGAGTTGGTA 4562
DB |||||
QY 4920 GGACAGTATTTGGTATCTCGGCTCTGCTGAAGCAGTAACTTCGGAAAGAGTTGGTA 4979
DB |||||
QY 4563 GCTCTTTGATCCGGCAAAACAAACACCGCTGGTATAGCGGTGTTTTTTTGGTAAAGAGTTCG 4622
DB |||||
QY 4980 GCTCTTTGATCCGGCAAAACAAACACCGCTGGTATAGCGGTGTTTTTTTGGTAAAGAGTTCG 5039
DB |||||
QY 4623 AGATTACGGCGCAGAAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGGGCTCG 4682
DB |||||
QY 5040 AGATTACGGCGCAGAAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGGGCTCG 5099
DB |||||
QY 4683 ACCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCAATGAGATTAATCAAAAAGGA 4742
DB |||||
QY 5100 ACCTCAGTGGAAACGAAACTCAGCTTAAGGGATTTTGGTCAATGAGATTAATCAAAAAGGA 5159
DB |||||
QY 4743 TCTTACCTAGATCTCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATAG 4802
DB |||||
QY 5160 TCTTACCTAGATCTCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAAAGTATATAG 5219
DB |||||
QY 4803 AGTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGCATCT 4862
DB |||||
QY 5220 AGTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGCATCT 5279
DB |||||
QY 4863 GTCTATTTCCGTTTCATCCATAGTTGCGCTGACTCCCGCTCGTGTAGATAAATACGATACGGG 4922
DB |||||
QY 5280 GTCTATTTCCGTTTCATCCATAGTTGCGCTGACTCCCGCTCGTGTAGATAAATACGATACGGG 5339
DB |||||
QY 4923 AGGGCTTACATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCACGCTACCGGCTC 4982
DB |||||
QY 5340 AGGGCTTACATCTGGCCCGCAGTGTGCAATGATACCGGAGACCCACGCTACCGGCTC 5399
DB |||||
QY 4983 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGCAGAGCGCAAGTGGTCTCGCAA 5042
DB |||||
QY 5400 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGCAGAGCGCAAGTGGTCTCGCAA 5459
DB |||||
QY 5043 CTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTGTTCGC 5102
DB |||||
QY 5460 CTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTGTTCGC 5519
DB |||||
QY 5103 CAGTTAATAGTTTGGCAACAGTTGTCGCATTCGCTACAGGCATCGTGGTGTCTACGCTCGT 5162
DB |||||
QY 5520 CAGTTAATAGTTTGGCAACAGTTGTCGCATTCGCTACAGGCATCGTGGTGTCTACGCTCGT 5579
DB |||||
QY 5163 CGTTTCGTATGGCTTCATTTACAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCC 5222
DB |||||
QY 5580 CGTTTCGTATGGCTTCATTTACAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCC 5639
DB |||||
QY 5223 CCATGTTGTCAAAAAGCGGTTAGTCTCTTCGGTCTCCCGATCGTTGTTCAGAAAGTAAGT 5282
DB |||||
QY 5640 CCATGTTGTCAAAAAGCGGTTAGTCTCTTCGGTCTCCCGATCGTTGTTCAGAAAGTAAGT 5699
DB |||||
QY 5283 TGGCCGAGTGTATACATCATCGGTTATGGCAGACTGCAATAATCTCTTACTGTGATGC 5342
DB |||||
QY 5700 TGGCCGAGTGTATCACTCATCGTGTATGGCAGACTGCAATAATCTCTTACTGTGATGC 5759
DB |||||
QY 5343 CATCCGTAAGATGCTTTTCTGCTGACTGGTGAAGTCAATCAACCAAGTCAATCTGAGAATAGT 5402

DB |||||
QY 5760 CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTCTCAACCAAGTCATTTCTGAGATAGT 5819
DB |||||
QY 5403 GTATCGGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGATATAACGGCGCCACATA 5462
DB |||||
QY 5820 GTATCGGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGATATAACGGCGCCACATA 5879
DB |||||
QY 5463 GCAGAACTTTAAAGTGCTCATCTTGGAAAGCTTTCTTCGGGGGCAAACTCTCAAGGA 5522
DB |||||
QY 5880 GCAGAACTTTAAAGTGCTCATCTTGGAAAGCTTTCTTCGGGGGCAAACTCTCAAGGA 5939
DB |||||
QY 5523 TCTTACCGCTGTGTAGATCCAGTTCCGATGTAACCCCACTCGTGACCCCAACTGATCTTCAG 5582
DB |||||
QY 5940 TCTTACCGCTGTGTAGATCCAGTTCCGATGTAACCCCACTCGTGACCCCAACTGATCTTCAG 5999
DB |||||
QY 5583 CATCTTTTACTTTCACAGCGTTTCTGGTGAGCAAAAACAGGAAGGCAAAATGCGCGCAA 5642
DB |||||
QY 6000 CATCTTTTACTTTCACAGCGTTTCTGGTGAGCAAAAACAGGAAGGCAAAATGCGCGCAA 6059
DB |||||
QY 5643 AAAGGGATAGGGGCGACACGGAATGTTGAATCTACTCATCTCTTCCCTTTTCAATATT 5702
DB |||||
QY 6060 AAAGGGATAGGGGCGACACGGAATGTTGAATCTACTCATCTCTTCCCTTTTCAATATT 6119
DB |||||
QY 5703 ATTGAAGCAATTTATCAGGGTTATTGTCTCATGCGGATACATATTTGAATGTATTAGA 5762
DB |||||
QY 6120 ATTGAAGCAATTTATCAGGGTTATTGTCTCATGCGGATACATATTTGAATGTATTAGA 6179
DB |||||
QY 5763 AAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCCACCTGACGTCTAAG 5822
DB |||||
QY 6180 AAAATAAACAAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCCACCTGACGTCTAAG 6239
DB |||||
QY 5823 AAACCAATTTATCATGACATTAACCTATAAAAATAGGCGTATCAGAGGCGCTTTTCGTC 5882
DB |||||
QY 6240 AAACCAATTTATCATGACATTAACCTATAAAAATAGGCGTATCAGAGGCGCTTTTCGTC 6299
DB |||||

RESULT 5
AAD06790 standard; DNA; 4276 BP.
ID AAD06790 standard; DNA; 4276 BP.
XX
AC AAD06790;
XX
DT 06-AUG-2001 (first entry)
XX
DE Plasmid pCMVII.
XX
KW Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;
KW HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;
KW plasmid pCMVII; cyclic; circular; ds.
XX
OS Human cytomegalovirus.
OS Homo sapiens.
OS Bos sp.
OS Unidentified.
OS Chimeric.
XX
PN WO200138358-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032249.
PF
XX 24-NOV-1999; 99US-0167224P.
PR
XX (CHIR) CHIRON CORP.
PA Selby M, Glazer E, Houghton M;
XX
XX WPI; 2001-367661/38.
DR
XX Virus-like particle for use as an immunogen, comprising a first hepatitis
PT B virus surface antigen (HBsAg) and chimeric antigen comprising a second
PT HBsAg covalently linked to hepatitis C immunogenic polypeptide.

XX	Disclosure; Fig 1; 115pp; English.	
XX	The invention relates to a virus-like particle for use as an immunogen,	
XX	comprising a first hepatitis B virus surface antigen (HBsAg) and a	
XX	chimeric antigen comprising a second HBsAg which is covalently linked to	
XX	an hepatitis C virus (HCV) immunogenic polypeptide, where the first and	
XX	the second HBsAg each comprise a substantially complete S domain. The	
XX	virus-like particle is useful as immunogen and as vaccine. The present	
XX	sequence is plasmid pCMVII, a pUC19-based cloning vector designed for	
XX	expression in mammalian cells. It comprises human cytomegalovirus (CMV)	
XX	immediate early (IE) enhancer/promoter, human CMV intron A, a human	
XX	tissue plasminogen activator (tPA) leader, a bovine growth hormone poly A	
XX	terminator (tSgt), a ColEI origin of replication and an ampicillin	
XX	resistance (Amp R) gene. This plasmid is used for expression of HBsAg and	
XX	chimeric antigen for production of virus-like particle of the invention	
XX	Sequence 4276 BP; 1043 A; 1086 C; 1067 G; 1080 T; 0 U; 0 Other;	
XX	Query Match 45.2%; Score 2658.4; DB 5; Length 4276;	
XX	Best Local Similarity 72.7%; Pred. No. 4.6e-276;	
XX	Matches 4275; Conservative 0; Mismatches 1; Indels 1606; Gaps 1;	
QY	1 TCGCGGTTTCGGTGATGACGGTGAACCTCTGCACATGCAGCTCCCGGAGACGGTCA 60	
DB	1 TCGCGCGTTTCGGTGATGACGGTGAACCTCTGCACATGCAGCTCCCGGAGACGGTCA 60	
QY	61 CAGCTTGTCTGAAGCGGATCGCGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGTG 120	
DB	61 CAGCTTGTCTGAAGCGGATCGCGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGTG 120	
QY	121 TTGGCGGGTGTGGGGCTGGCTTAACATATGCGGCATCAGACAGATTGTACTGAGAGTGC 180	
DB	121 TTGGCGGGTGTGGGGCTGGCTTAACATATGCGGCATCAGACAGATTGTACTGAGAGTGC 180	
QY	181 ACCATATGAAGCTTTTTCGAAAGCTTAGGCCCTCCAAAAGACCTCCTCCTACTTCTGG 240	
DB	181 ACCATATGAAGCTTTTTCGAAAGCTTAGGCCCTCCAAAAGACCTCCTCCTACTTCTGG 240	
QY	241 AATAGCTCAGAGCGCGAGCGGCTCGGCTCTGCTAATAAATAAATAAATAGTCAGGCCA 300	
DB	241 AATAGCTCAGAGCGCGAGCGGCTCGGCTCTGCTAATAAATAAATAAATAGTCAGGCCA 300	
QY	301 TGGGGCGGAGATGGGCGGAACTGGGCGGGAGGGAATATTGGCTATTGGCCCATTTGCAT 360	
DB	301 TGGGGCGGAGATGGGCGGAACTGGGCGGGAGGGAATATTGGCTATTGGCCCATTTGCAT 360	
QY	361 ACCTTGATCTATATCATTAATATGATCATTTATATGCGCTCATGTGCCAATGACGCCCA 420	
DB	361 ACCTTGATCTATATCATTAATATGATCATTTATATGCGCTCATGTGCCAATGACGCCCA 420	
QY	421 TGTTGACATTTGATTTGACTAGTTATTAATAGTAATCAATTTACGGGTCAATTAGTTTCAT 480	
DB	421 TGTTGACATTTGATTTGACTAGTTATTAATAGTAATCAATTTACGGGTCAATTAGTTTCAT 480	
QY	481 AGCCCATATATGGAGTTCCGCGGTACATAAATTTACGGTAAATGGCCCGCTGGCTGACCG 540	
DB	481 AGCCCATATATGGAGTTCCGCGGTACATAAATTTACGGTAAATGGCCCGCTGGCTGACCG 540	
QY	541 CCCAAACGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAACGCCAATA 600	
DB	541 CCCAAACGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAACGCCAATA 600	
QY	601 GGGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAATGCGCCACTTGGCAGTA 660	
DB	601 GGGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAATGCGCCACTTGGCAGTA 660	
QY	661 CATCAAGTGTATCATATGCCAAGTCCGCCCTTATTTGACGTCAATGACGGTAAATGGCCC 720	
DB	661 CATCAAGTGTATCATATGCCAAGTCCGCCCTTATTTGACGTCAATGACGGTAAATGGCCC 720	
QY	721 GCCTGGCATTTATGCCAGTACATGACCTTTACGGGACTTTTCCTACTTGGCAGTACATCTAC 780	
DB	721 GCCTGGCATTTATGCCAGTACATGACCTTTACGGGACTTTTCCTACTTGGCAGTACATCTAC 780	
QY	781 GTATTAGTCATCGCTATTACCATGGTGTATGCGGTTTTTGGCAGTACACCAATGGCGTGGGA 840	
DB	781 GTATTAGTCATCGCTATTACCATGGTGTATGCGGTTTTTGGCAGTACACCAATGGCGTGGGA 840	
QY	841 TAGCGGTTTGACTACACGGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900	
DB	841 TAGCGGTTTGACTACACGGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900	
QY	901 TTTTGGGACCAAAATCAACGGGACTTTTCCAAATGTGTGTAATAACCCGCCCGCTTGACG 960	
DB	901 TTTTGGGACCAAAATCAACGGGACTTTTCCAAATGTGTGTAATAACCCGCCCGCTTGACG 960	
QY	961 CAAATGGGCGGTAGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020	
DB	961 CAAATGGGCGGTAGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020	
QY	1021 CGTCAAGTCCCTTGAGAGCGGCATCCACGCTGTTTGGACCTCCATAGAAGCACCCGGGAC 1080	
DB	1021 CGTCAAGTCCCTTGAGAGCGGCATCCACGCTGTTTGGACCTCCATAGAAGCACCCGGGAC 1080	
QY	1081 CGATCCAGCTCCCGGCGCGGAAACGGTGCAATCGGAACCGGGAATCCCGCGTCCCAAGAGT 1140	
DB	1081 CGATCCAGCTCCCGGCGCGGAAACGGTGCAATCGGAACCGGGAATCCCGCGTCCCAAGAGT 1140	
QY	1141 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGGCTCTTATGTCATGCTATA 1200	
DB	1141 GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGGCTCTTATGTCATGCTATA 1200	
QY	1201 CTGTTTTGGCTTGGGGCTATACACCCCGCTCTTATGCTATAGTGATGGTATAGCT 1260	
DB	1201 CTGTTTTGGCTTGGGGCTATACACCCCGCTCTTATGCTATAGTGATGGTATAGCT 1260	
QY	1261 TAGCCTATAGGTGGGTTTATGACCACTTATGACCACTCCCTTATTTGGTGACCATACTT 1320	
DB	1261 TAGCCTATAGGTGGGTTTATGACCACTTATGACCACTCCCTTATTTGGTGACCATACTT 1320	
QY	1321 TCCATTACTAATCCATAACATGGGCTCTTTGGCACAACTATCTCTATTGGGCTATATGCAAA 1380	
DB	1321 TCCATTACTAATCCATAACATGGGCTCTTTGGCACAACTATCTCTATTGGGCTATATGCAAA 1380	
QY	1381 TACTCTGCTTTCAGAGACTGACACGGACTCTGTATTTTTTACAGGATGGGGTCCATTTAT 1440	
DB	1381 TACTCTGCTTTCAGAGACTGACACGGACTCTGTATTTTTTACAGGATGGGGTCCATTTAT 1440	
QY	1441 TATTTTACAAATTCACATATACAAACGCGCTCCCGTCCCGCCAGTCTTTTATTAAACA 1500	
DB	1441 TATTTTACAAATTCACATATACAAACGCGCTCCCGTCCCGCCAGTCTTTTATTAAACA 1500	
QY	1501 TAGCGTGGGATCTCCGACATCTCGGTACGTTTCCGAGCATGGGCTCTCTCCGGTAGC 1560	
DB	1501 TAGCGTGGGATCTCCGACATCTCGGTACGTTTCCGAGCATGGGCTCTCTCCGGTAGC 1560	
QY	1561 GCGGAGCTTCCACATCCGAGCCCTTGGTCCCATCCGTCAGCGGCTCATGTCGCTCGGC 1620	
DB	1561 GCGGAGCTTCCACATCCGAGCCCTTGGTCCCATCCGTCAGCGGCTCATGTCGCTCGGC 1620	
QY	1621 AGCTCTCTGCTCTTAAAGTGGAGGCCAGACTTAGGCACACAGCAATGCCACACACACC 1680	
DB		

QY	1861	AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCGGCCACACAGACATAATAGCT	1920
DB	1861		
QY	1861	AGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCGGCCACACAGACATAATAGCT	1920
DB	1861		
QY	1921	GACAGACTAACAGACTGTTCTTCCATGGTCTTTCTGCACTACCGTCGTCGACGAA	1980
DB	1921		
QY	1921	GACAGACTAACAGACTGTTCTTCCATGGGTCTTTCTGCACTACCGTCGTCGAC---	1977
DB	1921		
QY	1981	TTCAAGCAATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTGGAG	2040
DB	1978	-----	1977
QY	2041	CAGTCTTGTGTTGCCAGGCTAGCGAAACCCACGTCACCGGGGGAAGTGCGGGCCACA	2100
DB	1978	-----	1977
QY	2101	CTGTGCTGGATTGTTAGCTCTCGCACCGGCCAAGCAGACAGTCCAGCTGATCA	2160
DB	1978	-----	1977
QY	2161	ACACCAACGCGAGTTGGCACTCAATAGCACGGCCCTGAAC TGCAATGATGAGCTCAACA	2220
DB	1978	-----	1977
QY	2221	CCGGCTGTTGGCAGGCTTTTCTATCACACAAGTTCACTCTCAGGCTGTCCTGAGA	2280
DB	1978	-----	1977
QY	2281	GGCTAGCCAGCTGCCGACCCTTACCGATTTTGACCAGGCTGGGGCCCTATCAGTTATG	2340
DB	1978	-----	1977
QY	2341	CCAACGAAGCGGCCCGACAGCGCCCTACTGTGGCACTACCCGCCAAACCTTTGGC	2400
DB	1978	-----	1977
QY	2401	GTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTGCTTCACTCCAGCCCCGTGG	2460
DB	1978	-----	1977
QY	2461	TGGTGGGAACGCGACAGTCCGGCGCGCCACCTACAGCTGGGGTGAAATGATACGG	2520
DB	1978	-----	1977
QY	2521	ACGTCTTCGTCCTTAAACAATACCAGGCCACCGCTGGGCAATTGGTTCCGTTGTACTGGA	2580
DB	1978	-----	1977
QY	2581	TGAACCTCACTGGATTACCAAAGTGTGGAGGGCTCCTTGTGTATCGGAGGGCGG	2640
DB	1978	-----	1977
QY	2641	GCAACAACCCCTGCAGTCCCCACTGATTGCTTCCGCAAGCATCCGAGCCACATACT	2700
DB	1978	-----	1977
QY	2701	CTCGGTCCGGCTCCGGTCCCTGGATCACACCCAGGTCTGTGTCGACTACCCGTATAGGC	2760
DB	1978	-----	1977
QY	2761	TTTGGCATATCTTGTACCATCAACTACACCATAATTTAAATCAGGATGTACGTGGAG	2820
DB	1978	-----	1977
QY	2821	GGGTGGAACACAGCTGGAAGCTCCCTGCAACTGGACGGGGGGAACGTTGCGATCTGG	2880
DB	1978	-----	1977
QY	2881	AAGATAGGGAAGGTCCGAGTCGATATGGAGAACATCATCAGGATTCCTAGGACCCC	2940
DB	1978	-----	1977

QY	2941	TGCTCGTGTACAGGGGGGTTTTTCTTGTTCACAAGATCCTCACATAACCCAGAGTC	3000
DB	1978	-----	1977
QY	3001	TAGACTCGTGTGGACTTCTCTCAATTTTCTAGGGGATCTCCGTTGTCTTTGGCCAAA	3060
DB	1978	-----	1977
QY	3061	ATTGCGAGTCCCAACCTCCAATCATCTACCAACCTCTGTCTCTCOAATTTGTCTGGTT	3120
DB	1978	-----	1977
QY	3121	ATCGCTGGATGTCTGCGGCGTTTTATCATATTCTCTTCATCTGTCTGCTATGCCTCA	3180
DB	1978	-----	1977
QY	3181	TCTTCTTATTGTTCTTCTGGATTATCAAGTATGTTGCCCGTTTGTCTCTAATTTCCAG	3240
DB	1978	-----	1977
QY	3241	GATCAACAACACAGTACGGGACCATGCAAAACCTGCAAGACTCTCTGTCTCAAGGCAACT	3300
DB	1978	-----	1977
QY	3301	CTATGTTTCCCTCATGTTGCTGTACAAACCTACGGATGGAATTCACCTGTATTTCCCA	3360
DB	1978	-----	1977
QY	3361	TCCATCGTCTGGGCTTTCGCAAAATACCTATGGAGTGGGCTCAGTCCGTTCTCTT	3420
DB	1978	-----	1977
QY	3421	GGCTCAGTTTACTAGTGCATTTGTTCACTGTTTCTAGGGCTTTCCCCCACTGTTTGGC	3480
DB	1978	-----	1977
QY	3481	TTTCAGCTATATGGATGATGTGTTATTTGGGGCCCAAGTCTGTACAGATCGTGAGTCCCT	3540
DB	1978	-----	1977
QY	3541	TTATACCGCTGTTTCAATTTTCTTGTCTCTGGGTATACATTTTAAAGAAATTCAGACTCG	3600
DB	1978	-----CTAAGAAATTCAGACTCG	1994
QY	3601	AGCAAGTCTAGAAAGCGCGCCAGATATCAAGGATCCACTACGCTTATAGAGTCTCGTGA	3660
DB	1995	AGCAAGTCTAGAAAGCGCGCCAGATATCAAGGATCCACTACGCTTATAGAGTCTCGTGA	2054
QY	3661	TCAGCTCAGCTGCTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGGTGCCT	3720
DB	2055	TCAGCTCAGCTGCTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGGTGCCT	2114
QY	3721	TCTTGACCTCGAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGGAAATTTGCA	3780
DB	2115	TCTTGACCTCGAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGGAAATTTGCA	2174
QY	3781	TGCAATGTCTGAGTAGTGTCTATTCTATTCTGGGGGTGGGGTGGGGGAGGACAGCAAG	3840
DB	2175	TGCAATGTCTGAGTAGTGTCTATTCTATTCTGGGGGTGGGGTGGGGGAGGACAGCAAG	2234
QY	3841	GGGGAGATTGGGAAGACAATAGCAGGATGCTGGGAGCTCTTCCGCTTCTCGCTCAC	3900
DB	2235	GGGGAGATTGGGAAGACAATAGCAGGATGCTGGGAGCTCTTCCGCTTCTCGCTCAC	2294
QY	3901	TGACTCGCTGCGCTCGGTCTGGCTGGGGAGCGGTATCAGCTCACTCAAAAGCGGT	3960
DB	2295	TGACTCGCTGCGCTCGGTCTGGCTGGGGAGCGGTATCAGCTCACTCAAAAGCGGT	2354
QY	3961	AATACGGTTATCCACAGAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCCA	4020
DB	2355	AATACGGTTATCCACAGAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCCA	2414
QY	4021	GCAAAAGGCGAGAACCGTAAAAAGCGCGTGTGCTGGCGTTTTTTCATAGGCTCCGCC	4080

Db 2415 GCAAAAGCCAGGAACCGTAAAAAGCGCGGTGCTGGCGTGTTCATAGGCTCGCGCC 2474
QY 4081 CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACT 4140
Db 2475 CCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACT 2534
QY 4141 ATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCT 4200
Db 2535 ATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCT 2594
QY 4201 GCGCTTACCGGATACCTGTCGCGCTTTCCTCTTCGGAAGCGTGGCGCTTCTCTCAATG 4260
Db 2595 GCGCTTACCGGATACCTGTCGCGCTTTCCTCTTCGGAAGCGTGGCGCTTCTCTCAATG 2654
QY 4261 CTCACGCTGATGATCTCAGTTTCGCTGAGTGGTTCGCTCCCAAGCTGGGCTGTGCA 4320
Db 2655 CTCACGCTGATGATCTCAGTTTCGCTGAGTGGTTCGCTCCCAAGCTGGGCTGTGCA 2714
QY 4321 CGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAA 4380
Db 2715 CGAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAA 2774
QY 4381 CCCGGTAAAGACACACTTATCGCACCTGGCAGCAGCACCTGGTAAACAGGATTAAGCAGC 4440
Db 2775 CCCGGTAAAGACACACTTATCGCACCTGGCAGCAGCACCTGGTAAACAGGATTAAGCAGC 2834
QY 4441 GAGGTATGATAGCGGTGCTACAGTTCCTGAGTGGTGGCTTACCTTCGGAAGAGTTGG 4500
Db 2835 GAGGTATGATAGCGGTGCTACAGTTCCTGAGTGGTGGCTTACCTTCGGAAGAGTTGG 2894
QY 4501 AAGCACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGTTACCTTCGGAAGAGTTGG 4560
Db 2895 AAGCACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGTTACCTTCGGAAGAGTTGG 2954
QY 4561 TAGCTTTGATTCGGGCAAAACAAACACCGCTGCTGAGCGGTGGTTTTTTTTCGCAAGCA 4620
Db 2955 TAGCTTTGATTCGGGCAAAACAAACACCGCTGCTGAGCGGTGGTTTTTTTTCGCAAGCA 3014
QY 4621 GCAGTTACCGGCGAGAAAAGGATCTCAAGAGATCTTTCATCTTTCTACGGGCTC 4680
Db 3015 GCAGTTACCGGCGAGAAAAGGATCTCAAGAGATCTTTCATCTTTCTACGGGCTC 3074
QY 4681 TGACGCTCAGTGGAAAGAAACTCACGTTAAGGATTTTCGTCATGAGATTATCAAAAAG 4740
Db 3075 TGACGCTCAGTGGAAAGAAACTCACGTTAAGGATTTTCGTCATGAGATTATCAAAAAG 3134
QY 4741 GATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 4800
Db 3135 GATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATA 3194
QY 4801 TGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT 4860
Db 3195 TGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGAT 3254
QY 4861 CTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACG 4920
Db 3255 CTGCTATTTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACG 3314
QY 4921 GAGGGCTTACCTATCTGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGC 4980
Db 3315 GAGGGCTTACCTATCTGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGC 3374
QY 4981 TCCAGATTTATCAGCAATAAACCCAGCCGAGCGGAGCGGCGAGAGTGGTCTGTC 5040
Db 3375 TCCAGATTTATCAGCAATAAACCCAGCCGAGCGGAGCGGCGAGAGTGGTCTGTC 3434
QY 5041 AACTTTATCCGCTCCCATCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTTC 5100
Db 3435 AACTTTATCCGCTCCCATCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTTC 3494
QY 5101 GCCAGTTAATAGTTTGGCGCAACGTTGTGTCATTTGCTACAGGCATCGTGGTGTACGCTC 5160

Db 3495 GCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGTCTACAGGCATCGTGGTGTACGCTC 3554
QY 5161 GTCGTTTGTGATGGCTTTCATTAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATC 5220
Db 3555 GTCGTTTGTGATGGCTTTCATTAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATC 3614
QY 5221 CCCATGTTTGTGCAAAAAAGCGGTAGCTCTCTTCGGTCTCCGATCGTTGTCTAGAGTAA 5280
Db 3615 CCCATGTTTGTGCAAAAAAGCGGTAGCTCTCTTCGGTCTCCGATCGTTGTCTAGAGTAA 3674
QY 5281 GTTGGCGGAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATCTCTTACTGTCTCAT 5340
Db 3675 GTTGGCGGAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATCTCTTACTGTCTCAT 3734
QY 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCATTCTTGAGATA 5400
Db 3735 GCCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCATTCTTGAGATA 3794
QY 5401 GTGTATGCGGCGACCGAGTTGCTTTTCCCGCGCTCAATA CGGGATAATAACCGCGCACCA 5460
Db 3795 GTGTATGCGGCGACCGAGTTGCTTTTCCCGCGCTCAATA CGGGATAATAACCGCGCACCA 3854
QY 5461 TAGCAGAACTTTTAAAGTGTCTCATTTGGAACACGTTCTTCCGGGCGAAAACCTCTCAAG 5520
Db 3855 TAGCAGAACTTTTAAAGTGTCTCATTTGGAACACGTTCTTCCGGGCGAAAACCTCTCAAG 3914
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTCCAGTTTCGATGTAACCCACTCGTGCAACCACTGATCTTC 5580
Db 3915 GATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGCAACCACTGATCTTC 3974
QY 5581 AGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGAGCAAAAAAGGAGGCAAAAATGCCGC 5640
Db 3975 AGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGAGCAAAAAAGGAGGCAAAAATGCCGC 4034
QY 5641 AAAAAAGGAAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTTTTCAATA 5700
Db 4035 AAAAAAGGAAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTTTTCAATA 4094
QY 5701 TTATTCAAGCATTTATCAGGTTTATTGTTCTCATGAGCGGATACATATTTGAATGTATTTA 5760
Db 4095 TTATTCAAGCATTTATCAGGTTTATTGTTCTCATGAGCGGATACATATTTGAATGTATTTA 4154
QY 5761 GAAAAATAAAATAAGGGTTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGCTTA 5820
Db 4155 GAAAAATAAAATAAGGGTTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGCTTA 4214
QY 5821 AGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCAGAGGCCCTTTTCG 5880
Db 4215 AGAAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCAGAGGCCCTTTTCG 4274
QY 5881 TC 5882
Db 4275 TC 4276

RESULT 6

AD104102
ID AD104102 standard; DNA; 10263 BP.

XX AD104102;

XX AC AC

XX XX

XX 22-APR-2004 (first entry)

XX Vector pTnMod (CMV/Red).

DE

XX

XX

KW

KW

KW

KW

KW

KW

KW

KW

KW

KW

ss; vector; transposase; promoter; insertion sequence; Kozak; Tn10;
ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;
signal sequence; transgenic animal; Huntington's disease;
alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer;
cystic fibrosis; galactosaemia; congenital hypothyroidism;
maple syrup urine disease; neurofibromatosis; phenylketonuria;
sickle cell disease; Smith-Lemli-Optiz Syndrome; autoimmune disease;
shipping fever; cattle; mastitis; bacterial; viral; infection; Types I;

```

KW type; II; diabetes.
XX Synthetic.
XX Key Location/Qualifiers
XX misc_feature 1..130
FT /*tag= a
FT /note= "Remainder of F1(-) on from pBluescript11 sk(-)"
FT 131..132
FT /*tag= b
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT 133..1777
FT /*tag= c
FT /note= "CMV promoter/enhancer from pGwiz"
FT /note= "CMV promoter is modified by the addition of an
FT ACC sequence upstream of ATG"
FT 1778..1779
FT /*tag= d
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT 1780..2987
FT /*tag= e
FT /product= "Transposase"
FT /note= "Modified from Tn10 by optimising codons for
FT stability"
FT 2988..2993
FT /*tag= f
FT /note= "Two engineered stop codons"
FT 2994
FT /*tag= g
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT 2995..3410
FT /*tag= h
FT /note= "Taken from pGwiz"
FT 3415..3718
FT /*tag= i
FT /note= "Non-coding DNA residual from pNK2859"
FT 3719..3761
FT /*tag= j
FT /note= "Non-coding lambda DNA residual from pNK2859"
FT 3762..3831
FT /*tag= k
FT /note= "Left insertion sequence recognised by Tn10"
FT 3832..3837
FT /*tag= l
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT 3838..4044
FT /*tag= m
FT /note= "Multiple cloning site from pBluescript11 sk(-)"
FT /note= "May be used to insert the sequence of interest
FT into the vector"
FT 4049..5693
FT /*tag= n
FT /note= "CMV promoter/enhancer from pGwiz"
FT 5694..5701
FT /*tag= o
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT 5702..6617
FT /*tag= p
FT /note= "pRed reporter coding sequence including polyA
FT from pDsRed1.1"
FT 6618..7101
FT /*tag= q
FT /note= "Multiple cloning site from pBluescript11 sk(-)"
FT /note= "May be used to insert the sequence of interest
FT into the vector"
FT 7102..7106
FT /*tag= r
FT /note= "Residue from ligation of restriction enzyme sites

```

```

FT insertion_seq used in construction"
FT 7107..7176
FT /*tag= s
FT /note= "Right insertion sequence recognised by Tn10"
FT 7177..7218
FT misc_feature /*tag= t
FT /note= "Non-coding lambda DNA residual from pNK2859"
FT 7219..8062
FT misc_feature /*tag= u
FT /note= "Non-coding DNA residual from pNK2859"
FT 8063..10263
FT misc_feature /*tag= v
FT /note= "Remainder from pBluescript11 sk(-)"
FT
XX
XX WO2004003157-A2.
PN
PN 08-JAN-2004.
XX
XX 26-JUN-2003; 2003WO-US020389.
XX
XX 26-JUN-2002; 2002US-0392415P.
PR
PR 21-JAN-2003; 2003US-0441377P.
PR
PR 21-JAN-2003; 2003US-0441381P.
PR
PR 21-JAN-2003; 2003US-0441392P.
PR
PR 21-JAN-2003; 2003US-0441405P.
PR
PR 21-JAN-2003; 2003US-0441447P.
PR
PR 21-JAN-2003; 2003US-0441502P.
XX
XX (TRAN-) TRANSGENRES LLC.
PA
PA (LOUJ ) UNIV LOUISIANA STATE & AGRIC & MECH COLL.
XX
XX Cooper RK, Cadd GG, Fioretti WC, Deboer KF;
PI
PI WPI; 2004-083042/08.
XX
XX New vector comprising a transposase gene operably linked to a promoter,
PT useful in treating Alzheimer's disease, breast cancer, phenylketonuria,
PT autoimmune diseases, bacterial or viral infections or Types I and II
PT diabetes.
XX
XX Example 2; SEQ ID NO 2; 150pp; English.
XX
XX This sequence represents a new vector, pTnMod (CMV/Red), comprising a
CC transposase gene operably linked to a first promoter and one or more gene
CC of interest operably linked to one or more additional promoters, where
CC the one or more genes of interest and their operably linked promoters are
CC flanked by transposase insertion sequences recognized by the transposase
CC and the first promoter comprises a modified Kozak sequence comprising
CC ACCATG. The transposase is a Tn10 transposase and is modified in one to
CC twenty of the first codons. The gene of interest is operably linked to a
CC second promoter, which is a constitutive promoter or an inducible
CC promoter, e.g. an ovalbumin or a vitellogenin promoter. The vector
CC further comprises a polyA sequence, a conalbumin polyA sequence, operably
CC linked to the transposase gene. The vector further comprises two stop
CC codons operably linked to the transposase gene. The first or second gene
CC of interest is operably linked to a second or third promoter,
CC respectively. The first and a second gene of interest are operably linked
CC to a second promoter. The vector further comprises an enhancer operably
CC linked to the one or more genes of interest. The enhancer comprises at
CC least a portion of an ovalbumin enhancer. The vector further comprises an
CC egg directing sequence, ovalbumin or ovomucoid signal sequence or
CC vitellogenin targeting sequence, operably linked to the one or more genes
CC of interest. The vector is useful in producing transgenic animals with
CC desired proteins or molecules. They are also useful in treating
CC Huntington's disease, alpha-1-antitrypsin deficiency, Alzheimer's
CC disease, breast cancer, cystic fibrosis, galactosaemia, congenital
CC hypothyroidism, maple syrup urine disease, neurofibromatosis,
CC phenylketonuria, sickle cell disease, Smith-Lemli-Optiz Syndrome,
CC autoimmune diseases, shipping fever in cattle, mastitis, bacterial or
CC viral infections or Types I and II diabetes.
XX
XX Sequence 10263 BP; 2648 A; 2595 C; 2374 G; 2646 T; 0 U; 0 Other;
SQ

```


Db 6156 GGCTCCACCGAGCGCTGTACCCCGCGAGCGGTGCTGAAGGCGAGATCCACAAGGC 6215
QY 2357 CGACGACGCGCCCTACTGCTGGCACTACC-----CCCCAA 2391
Db 6216 CCTGAAGCTGAAGGACGGCGCCACTACTCTGGTGGAGTTCAAAGTCCATCTACATGGCCAA 6275
QY 2392 AACCTTGGCGTATGTGCCCGGAAGAGTGTGTGGTCCGGTATATTTGCTTCACTCC- 2450
Db 6276 GAAGCCGCTGCAGTCCCGGCTACTACTACGTGGACTCCAACTGGACATCACTCCCA 6335
QY 2451 -----AGCCCGTGTGGTGGGAACGACCGACAGGTGGCGCGGCCACCTACA 2499
Db 6336 CAACGAGGACTACCACTGTGAGCAGTAGAGGCCACCGAGGGCGCCACCCTGTT 6395
QY 2500 GCTGGGTGAAAAATGA-----TACGGAGCTTCTCGT 2530
Db 6396 CCTGTAGGCGCGGACTCTAGATCATATAATCAGCCATACCACATTTGTAGAGGTTTACT 6455
QY 2531 CCTTAACAATACAGGCGCCGCTGGCAATTGGTTCGGTTGTACTGGATGACTCAAC 2590
Db 6456 TGCTTTAAAAAACCTCCACACCTCCCTGAAACCTGAAACATAAAATGAATGCAATTGT 6515
QY 2591 TGGATTACCAAAAGTGTGGGAGCGCTCCTGTGTATCGGAGGGCGGCGCAACAACAC 2650
Db 6516 TGTGTAACTTGTATTGACGCTTATAATGTTACAAATAAAGCAATAGCATCAAA 6575
QY 2651 CCT-----GCACTGCCCCACTGATTC 2672
Db 6576 TTTCACAAATAAAGCATTTTTTCACTGCATCTAGTTGTGGCCCGGCTGCAGGAATTC 6635
QY 2673 TTCGGAAGATCCGAGCGCACATCTCTCGTGGCGTCCGGTCCCTGGATCACAACC 2732
Db 6636 GATATCAAGCTTATCGATACCGGTGACCTCGAGGGGGGCGCGGTACCCAAATTCGCCCTA 6695
QY 2733 AGTGTCTGTCTGACTACCG--TATAGGCTTTGGCATTCCTGTACCATCAACTACA 2790
Db 6696 TAGTGAGTCGTATTACGGCGCTCACTGGCCGTCGTTTTACAAGTCGTGACTGGGAATA 6755
QY 2791 CCATATTTAAATCAGGATGTAGTGGAGGGTTCGAACACAGGCTGGAAGCTGCCTGCA 2850
Db 6756 CCCTGGGTTACCACTTAATCGCTTGCGACACATCCCTTTTCGCGAGCTGGCGTAA 6815
QY 2851 ACTGGACGGGGCGGACGTTGGATCTGGAAGATAGGACAGGTC----- 2897
Db 6816 TAGCGAAGAGCGCCGACCGATCGCCCTTCCCAACAGTTTGGCGAGCTGGAATGCGAATG 6875
QY 2898 GAGATCGATATGAGAACATCACAATCAGGATCTTAGGACCCCTGCTCGTGTACAGGCG 2957
Db 6876 GAAATTTGAAGCGTTAAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCA 6935
QY 2958 GGTTTTTTCTTTGTGAAGAATCCTCACAATACCG----- 2993
Db 6936 TTTTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAG 6995
QY 2994 -----CAGAGTCTAGACTCGTGGTGACTTCTCTC 3023
Db 6996 ATAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAGAACAAGTGGACTCC 7055
QY 3024 AATTTTCTAGGGGATCTCCCGTGTGCTTGGCA----- 3058
Db 7056 AACGTCAAGGGCGAAACCGTCTATCAGGGCGATGGCCCACTAATCCGGGATCATATG 7115
QY 3059 ----- 3058
Db 7116 ACAAGATGTATCCACCTTAACTTAATGATTTTTTACAAAATCAATTAGGGGATTCATCA 7175
QY 3059 ----- 3058
Db 7176 GTGCTCAGGGTCAACGAGAAATTAACATTCGTCAGGAAGCTTATGATGATGATGTGCTT 7235
QY 3059 -----AAATTCGAGTCCCAACCTCCAACTCACTCACAACCTC 3097
Db 7236 AAAAACTTACTCAATGGCTGTTATGCATATCGCATATCGAATACATCGAATAACCTTAAAGAGC 7295

QY 3098 CTGTCTCCOATTTGTCTGGTTATTCGCTGGATGTGTCTTGGCGGCTTT----- 3145
Db 7296 TTGCCGATAAAAAAGGCCAAATTTATGTCTATTTACCGCGGCTTTTATTTAGCTGTAAG 7355
QY 3146 -----TATCATATTCCTCTTCACTCTGCTGCTATGCTCATCTTCTT 3187
Db 7356 ATAAATAAATAAGATAGGTTTTTATTTGAAGCTAAATCTTCTTATCGTAAAAAATGCCCT 7415
QY 3188 ATTTGTTCTTCTGGATTATCAAGGTATGTTGCCG----- 3222
Db 7416 CTTGGGTTATCAAGGGTCAATATATTTTCGCGGAATAACATCAATTTGGTGACAAATAA 7475
QY 3223 ----- 3222
Db 7476 CTAAGCACTGTCTCTCTGTTTACTCCCTGAGCTTGAGGGGTAAACATGAAGGTATCA 7535
QY 3223 -----T 3223
Db 7536 TAGCAGGATAAATAACAGTAATAACGCTAAACCAATAATCCAAATCCAGCCATCCCAAT 7595
QY 3224 TTGTCTCTTAATTTCCAGGATCAACAACCAACAGTACGGGACCATGCAAAACCTGACGAC 3283
Db 7596 TGGTAGTGAATGATTATAATAACAGCAACAGTAATGGCCAAATAACACCGTTCATT 7655
QY 3284 TCTGTCTCAAGGAACTCTATGTTTCCCTCATGTTGCTG----- 3322
Db 7656 GGTAAAGGCTCACCAATAATCCCTGTAAAGACCTGCTGATGACTCTTTGTTTGGATAGA 7715
QY 3323 -----TACAAAACCTACGGATGGAATTTGCACCTG 3352
Db 7716 CATCACTCCTGTAAATGCAAGTAAAGGATCCCAACAGCCCAATAAATTTAAACAAG 7775
QY 3353 TATTCCTCATCCCATCGTCTGGGCTTTTCGCAAAATACCTATGGGAGTGGGCCCTCAGTCCG 3412
Db 7776 GAAACTAACCACCTTCAGATATAAACGCTAAAGGCAATGCACTACTATCTGCAAT 7835
QY 3413 TTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTGAGTGTGCTGAGGCTTTCCCCAC 3472
Db 7836 AAATCCGAGCAGTACTGCGGCTTTTTCGCCCATTTAGTGGCTATTTCTCTGTCACAAAG 7895
QY 3473 TGTGTTGGCTTTTCAGCTATA-----TGATGATGTGTTATTTGGGGCCCAAG-----TCTGT 3522
Db 7896 GCTTGAATFACTGAGTGTAAAGACCAAGACCGGTAATGAAAGCCCAACCATCATCTAT 7955
QY 3523 ACAGCATGTGAGTCCCTTTATACCGCTGTTACCAATTTTCTTTGCTCTGGGTATACA 3582
Db 7956 TCATCATCAGATTTCTGTAATAGCACACACCGTGTGGATTGGCTATCAATGCGCTGA 8015
QY 3583 TTTAAGAAATTCAGATCTGAGCAAGTCTAGAAGGCGGCCAAGATATCAAGATCCACTA 3642
Db 8016 AATAATAATCAACAAATGGCATCGTTAAATAAGTATGATATACCGATCAGCTTTTGTTC 8075
QY 3643 CGGTTTAGAGCTCGCTGATCAGCTCGAC-----TGTGCTTCTAGTTGCCAGCATC 3695
Db 8076 CTTTAGTAGGGTTAATTCGCGCTTGGCGTAAATCATGTCATAGCTGTTTCTGTGTGA 8135
QY 3696 TGTGTTTTCCTCCCGTGCCTTTCTTGACCCCTGGAAGGTGCCATCCCACTGTCTCT 3755
Db 8136 AATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCC 8195
QY 3756 TTCCTAATAAATAGGAAATTCATCGCATGCTGATGCTGAGTAGTGTCTATTTCTGTTGG 3815
Db 8196 TGGGTTGCCATAAGTAGTGAGCTAACTCAATTAATTTGCGTTGCGCTCACTGCCCGCTTC 8255
QY 3816 GGTGGGG--TGGGCGAGACAGCAAGGGGAGGATTCGGAAGACAATAGACAGGATGC 3872
Db 8256 CAGTCGGGAAACCTGTCTGTCGCGCTGCAATTAATGAATCGGCCAACGCGGGGAGAGGC 8315
QY 3873 -----TGGGAGCTTTCCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTT 3922
Db 8316 GGTTCGCTATTGGGGCTCTTCGCTTCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTT 8375

Qy	3923	CGGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCA	3983
Db	8376	CGGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCA	8435
Qy	3983	GGGGATAACGACAGGAAGAAACAATGTGAGCAAAAGCCAGCAAAAAGGCCACGAGACCGTAAA	4042
Db	8436	GGGGATAACGACAGGAAGAAACAATGTGAGCAAAAGGCCAGCAAAAAGGCCAGGACCGTAAA	8495
Qy	4043	AAGGCGCGTGTGCTGGCGTTTTTCCATATGAGCTCCGCCCCCTGACGAGCATCAAAAAAT	4102
Db	8496	AAGGCGCGTGTGCTGGCGTTTTTCCATATGAGCTCCGCCCCCTGACGAGCATCAAAAAAT	8555
Qy	4103	CGACGCTCAAGTTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCC	4162
Db	8556	CGACGCTCAAGTTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCC	8615
Qy	4163	CCTGGAGCTCCCTCGTCGCTCTCTGTTCCGACCTGCGCGTTTACCGGATACCTGTCC	4222
Db	8616	CCTGGAGCTCCCTCGTCGCTCTCTGTTCCGACCTGCGCGTTTACCGGATACCTGTCC	8675
Qy	4223	GCCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCAATGCTCACGCTCTGAGTATCTCAGT	4282
Db	8676	GCCTTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCAATGCTCACGCTCTGAGTATCTCAGT	8735
Qy	4283	TCGGTGTAGTTCGTTTCGTTCCAAAGCTGGCTGTGTGACGAAACCCCGCTTACGCCGAC	4342
Db	8736	TCGGTGTAGTTCGTTTCGTTCCAAAGCTGGCTGTGTGACGAAACCCCGCTTACGCCGAC	8795
Qy	4343	CGCTCGCGCTTATCCGCTTAACTATCTGCTTGGTTCGAGTTCGAGTTCGAGTTCG	4402
Db	8796	CGCTCGCGCTTATCCGCTTAACTATCTGCTTGGTTCGAGTTCGAGTTCGAGTTCG	8855
Qy	4403	CCACTGGCAGCAGCCACTGGTTAAACAGGATAGCAGAGCGAGGTATGTAGCGGCTGTACA	4462
Db	8856	CCACTGGCAGCAGCCACTGGTTAAACAGGATAGCAGAGCGAGGTATGTAGCGGCTGTACA	8915
Qy	4463	GAGTTCTTGAAGTGGTGGCTTAACTACGGCTTACCTAGAGGACAGTATTTGGTATCTGC	4522
Db	8916	GAGTTCTTGAAGTGGTGGCTTAACTAGAGGACAGTATTTGGTATCTGC	8975
Qy	4523	GCTCTGCTGAAGCAGTATACCTTCGGAAAGAGTTCGTAGCTCTTCGATCCGCGAACAA	4582
Db	8976	GCTCTGCTGAAGCAGTATACCTTCGGAAAGAGTTCGTAGCTCTTCGATCCGCGAACAA	9035
Qy	4583	ACCACCGCTGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACCGCGACGAAAAA	4642
Db	9036	ACCACCGCTGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACCGCGACGAAAAA	9095
Qy	4643	GGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAC	4702
Db	9096	GGATCTCAAGAAGATCTTTTGATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAAC	9155
Qy	4703	TCAGTTTAAGGATTTTGGTCATCAGATTTATCAAAAAGGATCTTCACCTAGATCTTTTA	4762
Db	9156	TCAGTTTAAGGATTTTGGTCATCAGATTTATCAAAAAGGATCTTCACCTAGATCTTTTA	9215
Qy	4763	AAATTAATAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACGT	4822
Db	9216	AAATTAATAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACGT	9275
Qy	4823	TACCAATGCTTAAATCAGTGGACCACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATA	4882
Db	9276	TACCAATGCTTAAATCAGTGGACCACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATA	9335
Qy	4883	GTTGCGCTGACTCCCGCTCGTGTAGATAAATCTACGATACGGGAGGCTTTACCATCTGGCCCC	4942
Db	9336	GTTGCGCTGACTCCCGCTCGTGTAGATAAATCTACGATACGGGAGGCTTTACCATCTGGCCCC	9395
Qy	4943	AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAAC	5002
Db	9396	AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATAAAC	9455
Qy	5003	CAGCCAGCCGGAAGGGCGGACGAGGAGTGGTTCCTCAACTTTTATCCGCTCCATCCAG	5062

Db	9456	CAGCCAGCCGGAAAGGCGGAGAGTGGCTCGTCAACTTTATTCCGCCTCATCCAG	9515
Qy	5063	TCATTAAATGTTGCGCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTCGCGCAAC	5122
Db	9516	TCATTAAATGTTGCGCGGAAGCTAGNGTAAGTAGTTCGCCAGTTAATAGTTCGCGCAAC	9575
Qy	5123	GTTGTTGCCAATGCTACAGGCATCTGGTGTCACGCTGTCGTGTTGGTATGGCTTCATTC	5182
Db	9576	GTTGTTGCCAATGCTACAGGCATCTGGTGTCACGCTGTCGTGTTGGTATGGCTTCATTC	9635
Qy	5183	AGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAGCG	5242
Db	9636	AGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAGCG	9695
Qy	5243	GTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAAGTAAGTTGGCCGAGTGTTCATCACTC	5302
Db	9696	GTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAAGTAAGTTGGCCGAGTGTTCATCACTC	9755
Qy	5303	ATGGTTATGCGACACTGCATAAATCTCTTTCATGTCATGCCATCCGTAAGATGCTTTTCT	5362
Db	9756	ATGGTTATGCGACACTGCATAAATCTCTTTCATGTCATGCCATCCGTAAGATGCTTTTCT	9815
Qy	5363	GTGACTGGTAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTTGC	5422
Db	9816	GTGACTGGTAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGCGACCGAGTTGC	9875
Qy	5423	TCTTGCGCGCGCTCAATACGCGGATAATACCGCGCCACATAGCAGAACTTTTAAAGTGCTC	5482
Db	9876	TCTTGCGCGCGCTCAATACGCGGATAATACCGCGCCACATAGCAGAACTTTTAAAGTGCTC	9935
Qy	5483	ATCAATTGGAAAAAGTTCCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCC	5542
Db	9936	ATCAATTGGAAAAAGTTCCTTCGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCC	9995
Qy	5543	AGTTCGATGTAAACCACTCGTGACACCAACTGATCTTCAGCATCTTTACTTTCCACCAGC	5602
Db	9996	AGTTCGATGTAAACCACTCGTGACACCAACTGATCTTCAGCATCTTTACTTTCCACCAGC	10055
Qy	5603	GTTTCTGGGTGACAAAAACAGGAAGCAAAATGCCGCAAAAAGGAATTAAGGCGGACA	5662
Db	10056	GTTTCTGGGTGACAAAAACAGGAAGCAAAATGCCGCAAAAAGGAATTAAGGCGGACA	10111
Qy	5663	CGGAAATGTTGGAATACTCATACTCTTCCTTTTCAATATTTATTTGAAGCATTTATCAGGGT	5722
Db	10116	CGGAAATGTTGGAATACTCATACTCTTCCTTTTCAATATTTATTTGAAGCATTTATCAGGGT	10171
Qy	5723	TATTGCTCATGACGGATACATATTTGAATGTATTTAGAAAAATTAACAATAGGGGTT	5782
Db	10176	TATTGCTCATGACGGATACATATTTGAATGTATTTAGAAAAATTAACAATAGGGGTT	10233
Qy	5783	CCGCGCACATTTCCCGAAAAAGTGCCAC	5810
Db	10236	CCGCGCACATTTCCCGAAAAAGTGCCAC	10263
RESULT 7			
AAF83668			
ID	AAF83668 standard; DNA; 4282 BP.		
XX	AAF83668;		
XX	23-JUL-2001 (first entry)		
DT	pCMV-II nucleic acid sequence.		
DE	HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;		
XX	catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;		
KW	immunotherapy; NS3; ds.		
KW	Synthetic.		
OS	Human cytomegalovirus.		
OS	Bos sp.		

XX	W0200138360-A2.	QY	481	AGCCCATATATGGAGTTCCGCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCG	540
PN	31-MAY-2001.	DB	481	AGCCCATATATGGAGTTCCGCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCG	540
XX	22-NOV-2000; 2000WO-US032326.	QY	541	CCCAACGACCCCGCCCAATTGACGTCAATAATGACGTATGTTCCTCATAGTAAGCCCAATA	600
PD	24-NOV-1999; 99US-0167502P.	DB	541	CCCAACGACCCCGCCCAATTGACGTCAATAATGACGTATGTTCCTCATAGTAAGCCCAATA	600
XX	(CHIR) CHIRON CORP.	QY	601	GGGACTTTCATTTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	660
XX	Coit D, Medina-Selby A, Selby M, Houghton M;	DB	601	GGGACTTTCATTTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	660
XX	WPI; 2001-343948/36.	QY	661	CATCAAGTGTATCATATGCGCAAGTCGCGCCCTATTGACGTCAATGACCGTAAATGSCCC	720
DR	Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as	DB	661	CATCAAGTGTATCATATGCGCAAGTCGCGCCCTATTGACGTCAATGACCGTAAATGSCCC	720
XX	a vaccine against HCV, comprises a polypeptide having a mutation that	QY	721	GCTGGCATTTATGCCCCAGTACATGACCTTACCGGACTTTCCTACTTGGCAGTACATCTAC	780
PT	functionally disrupts the catalytic domain of NS3.	DB	721	GCTGGCATTTATGCCCCAGTACATGACCTTACCGGACTTTCCTACTTGGCAGTACATCTAC	780
XX	Example 1; Fig 7; 340pp; English.	QY	781	GTATTAGTCAATCGCTATTACCATGATGCGGTTTGGCAGTACACCAATGGCGGTGA	840
XX	The invention relates to an isolated mutant non-structural (NS) Hepatitis	DB	781	GTATTAGTCAATCGCTATTACCATGATGCGGTTTGGCAGTACACCAATGGCGGTGA	840
CC	C virus (HCV) polypeptide, comprising a polypeptide having a mutation in	QY	841	TAGCGGTTTGAAGTCTACGGGGATTTCCAACTCTCCACCCCATTTGACGTCAATGGGAGTTTG	900
CC	the catalytic domain of NS3, where the mutation functionally disrupts the	DB	841	TAGCGGTTTGAAGTCTACGGGGATTTCCAACTCTCCACCCCATTTGACGTCAATGGGAGTTTG	900
CC	catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a	QY	901	TTTGGGACCAAAAATCAACGGGACTTTCCAAATAATGTCGTAATAACCCGCCCGTTGACG	960
CC	and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide	DB	901	TTTGGGACCAAAAATCAACGGGACTTTCCAAATAATGTCGTAATAACCCGCCCGTTGACG	960
CC	in vaccines against HCV and as diagnostics. The antibodies raised against	QY	961	CAATGGGCGGTAGGCGTGTAGCGTGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC	1020
CC	these polypeptides can also be used as diagnostics, or for passive	DB	961	CAATGGGCGGTAGGCGTGTAGCGTGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC	1020
CC	immunotherapy. The antibodies are also useful for isolating and	QY	1021	CGTCAGATCGCTGGAGAGCCCATCACGCTGTTTTGACCTCCATAGAGACACCCGGAC	1080
CC	identifying HCV particles. The present sequence represents the nucleic	DB	1021	CGTCAGATCGCTGGAGAGCCCATCACGCTGTTTTGACCTCCATAGAGACACCCGGAC	1080
CC	acid sequence of the pCMV-11 containing the human CMV promoter, enhancer,	QY	1081	CGATCCAGGCTCCGCGCGCGGAAACGGTGCATTTGGAACCGGATTTCCCGTGCACAGAGT	1140
CC	intron A, polylinker and the bovine growth hormone terminator in a	DB	1081	CGATCCAGGCTCCGCGCGCGGAAACGGTGCATTTGGAACCGGATTTCCCGTGCACAGAGT	1140
CC	deleted-pUC backbone	QY	1141	GACGTAAGTACCGCTATAGACTCTATAGGACACACCCCTTTGGCTCTTATGCAATGCTATA	1200
SQ	Sequence 4282 BP; 1044 A; 1089 C; 1068 G; 1081 T; 0 U; 0 Other;	DB	1141	GACGTAAGTACCGCTATAGACTCTATAGGACACACCCCTTTGGCTCTTATGCAATGCTATA	1200
	Query Match 44.2%; Score 2601.2; DB 4; Length 4282;	QY	1201	CTGTTTTTGGCTTTGGGGCTTATACACCCCGC-TCCCTATGCTATAGGTGATGCTATAGC	1259
	Best Local Similarity 72.5%; Pred. No. 5.8e-270;	DB	1201	CTGTTTTTGGCTTTGGGGCTTATACACCCCGC-TCCCTATGCTATAGGTGATGCTATAGC	1259
	Matches 4269; Conservative 0; Mismatches 80; Indels 1612; Gaps 5;	QY	1260	TTAGCCCTATAGGTGTGGGTATTGACCAATTTGACCACTCCCTTATTTGGTGAAGTACT	1319
QY	1 TCGCGGTTTCGTTGATGACGGTGAAACCTCTGACACATGACAGTCCCGGAGACGGTCA	DB	1260	TTAGCCCTATAGGTGTGGGTATTGACCAATTTGACCACTCCCTTATTTGGTGAAGTACT	1319
DB	1 TCGCGGTTTCGTTGATGACGGTGAAACCTCTGACACATGACAGTCCCGGAGACGGTCA	QY	1320	TTCCATTAATCAATCAATGAGTCTTTGGCAAACTATCTCTATTGGCTATATGCGCA	1379
QY	61 CAGCTTTGCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCAAGCGGTG	DB	1320	TTCCATTAATCAATCAATGAGTCTTTGGCAAACTATCTCTATTGGCTATATGCGCA	1379
DB	61 CAGCTTTGCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCAAGCGGTG	QY	1380	TTCCATTAATCAATCAATGAGTCTTTGGCAAACTATCTCTATTGGCTATATGCGCA	1380
QY	121 TTGGCGGTTTCGCGGCTGCTTAACCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC	DB	1380	TTCCATTAATCAATCAATGAGTCTTTGGCAAACTATCTCTATTGGCTATATGCGCA	1380
DB	121 TTGGCGGTTTCGCGGCTGCTTAACCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC	QY	1381	ATACTCTGCTTTCAGAGACTGACCGGACTCTGTTATTTTACAGATGGGTCCCATTT	1438
QY	181 ACCATATGAAGCTTTTGGAAAAGCCTTAGCCCTCCAAAAGCCTCTCTACTACTTCTGG	DB	1381	ATACTCTGCTTTCAGAGACTGACCGGACTCTGTTATTTTACAGATGGGTCCCATTT	1438
DB	181 ACCATATGAAGCTTTTGGAAAAGCCTTAGCCCTCCAAAAGCCTCTCTACTACTTCTGG	QY	1439	ATTATTTACAAATTTACATATACAAACGCGCGTCCCGCTGCGGAGTGGGTATTTAA	1498
QY	241 AATAGCTCAGAGCCGAGCGGCTCGGCCCTCTGCATAAATAAAAAAATTTAGTCAGCCA	DB	1439	ATTATTTACAAATTTACATATACAAACGCGCGTCCCGCTGCGGAGTGGGTATTTAA	1498
DB	241 AATAGCTCAGAGCCGAGCGGCTCGGCCCTCTGCATAAATAAAAAAATTTAGTCAGCCA	QY	1441	ATTATTTACAAATTTACATATACAAACGCGCGTCCCGCTGCGGAGTGGGTATTTAA	1500
QY	301 TGGGCGGAGAAATGGCGGAACTGGCGGGGAGGAAATTTGGCTATTGGCCATTGCAT	DB	1441	ATTATTTACAAATTTACATATACAAACGCGCGTCCCGCTGCGGAGTGGGTATTTAA	1500
DB	301 TGGGCGGAGAAATGGCGGAACTGGCGGGGAGGAAATTTGGCTATTGGCCATTGCAT	QY	1499	CATAGCGTGGGATCTCC---GACATCTCGGTACGTTTCCGAGCATGGGCTCTTCCG	1555
QY	361 ACCTTGATCTATCATATATATGATCATTTTATTTGGCTCATGTCGAATATGACGCCA	DB	1499	CATAGCGTGGGATCTCC---GACATCTCGGTACGTTTCCGAGCATGGGCTCTTCCG	1555
DB	361 ACCTTGATCTATCATATATATGATCATTTTATTTGGCTCATGTCGAATATGACGCCA	QY	1501	CATAGCGTGGGATCTCCAGCGGAATCTCGGGTACGTTTCCGAGCATGGGCTCTTCCG	1560
QY	421 TGTGTGACATTGATTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTCAT	DB	1501	CATAGCGTGGGATCTCCAGCGGAATCTCGGGTACGTTTCCGAGCATGGGCTCTTCCG	1560
DB	421 TGTGTGACATTGATTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTCAT	QY	1556	GTAGCGGCGGAGCTTCCACATCCGAGCCCTGGTGGTCCCATCCGTCCAGCGGCTCATGTGCG	1615

1561 GTAGCGGAGGCTTCACATCCGAGCCCTGGTCCCATCGCTCCAGCGGCTCATGTGCG 1620
1616 TCGGAGCTCCTTGTCTTAACTAGTGGAGCCAGACTTAGGCGACAGACACATGCCACCA 1675
1621 TCGGAGCTCCTTGTCTTAACTAGTGGAGCCAGACTTAGGCGACAGACACATGCCACCA 1680
1676 CCACAGTGTGCGGCAAGCGCGTGGCGGTAGGTATGTCTGAAATGAGCTCGGAG 1735
1681 CCACAGTGTGCGGCAAGCGCGTGGCGGTAGGTATGTCTGAAATGAGCTCGGAG 1740
1736 ATTGGGCTCCAGC - TGGAGCGAGATGGAAGACTTAAAGGCGAGCGGCGAGAGAGATGCGAG 1794
1741 ATTGGGCTCCAGCCTGAGCGAGATGGAAGACTTAAAGGCGAGCGGCGAGAGAGATGCGAG 1800
1795 GCAGCTGAGTGTGTGTTATCTGATATAGAGTACAGGTAACTCCCGTTGCGGTGCTGTTAA 1854
1801 GCAGCTGAGTGTGTGTTATCTGATATAGAGTACAGGTAACTCCCGTTGCGGTGCTGTTAA 1860
1855 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGGCGCGGCCACACAGACATA 1914
1861 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCGGCGCGGCCACACAGACATA 1920
1915 ATAGCTGACAGACTAAGACTGTTCTTCCATGGGTCTTTTCTGAGTACAGTACCGTGGTC 1974
1921 ATAGCTGACAGACTAAGACTGTTCTTCCATGGGTCTTTTCTGAGTACAGTACCGTGGTC 1980
1975 GACGAATTCAGCAATCATGATGATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGT 2034
1981 GAC----- 1983
2035 GTGAGCAGTCTTCTGTTTCGCCAGCGCTAGCGAAACCCACGTCACGGGGGAAGTGCCG 2094
1984----- 1983
2095 GCCACACTGTGTGATTTGTAGCTCTCTCGACACAGGCGCAAGAGAACGTCCAGC 2154
1984----- 1983
2155 TGATCAACACCAAGCGAGTTGGACCTCAATAGCAGCGCCCTGAACGTGAATGATAGCC 2214
1984----- 1983
2215 TCAACACCGCTGTTGGCAGGCGTTTCTATCACCACAAAGTTCAACTCTTCAGGCTGTC 2274
1984----- 1983
2275 CTGAGAGGCTAGCCAGTCCGACCCCTTACCGATTTTGAACAGGCTGGGGCCCTATCA 2334
1984----- 1983
2335 GTTATGCCAAGGAGCGGCCCGACGAGCGCCCTACTGCTGGCACTACCCGCCAAAC 2394
1984----- 1983
2395 CTTGCGGTATGTGCGCGGAGAGTGTGTGCTCGGTATATGTCTTCACTCCAGCC 2454
1984----- 1983
2455 CCGTGTGTGGAAACGACCGACAGTTCGGGCGCGCCCACTACAGCTGGGGTGAAATG 2514
1984----- 1983
2515 ATACGAGCGTCTTCGTCTTAAACATACAGGCCACCGCTGGGCAATTGGTTCCGTTGTA 2574
1984----- 1983
2575 CCTGGATGAACCTCACTGGATTACCAAAAGTGTGCGGAGCGCTCTTGTGTATCGGAG 2634
1984----- 1983
2635 GGGCGGCAACAAACCTTGCACTGGCCCACTGATTGCTTCCGAAAGCATTCGGACGCCA 2694

1984----- 1983
2695 CATACTCTCGGTGCGGCTCGGCTCCCTGGATCACACCAGGTGCTGGTGCAGTACCCGT 2754
1984----- 1983
2755 ATAGGCTTTTGGCATTAATCTTGTATCATCAACTACACCATATTTTAAATCAGGATGTACG 2814
1984----- 1983
2815 TGGAGGGGTGCAACACAGCGCTGGAAGCTGCCTGCAACTGGAGCGGGGGGGAAGTTGCG 2874
1984----- 1983
2875 ATCTGGAAGATAGGACAGGTCCGAGATCGATATGGAAGACATCACATCAGGATTCCTAG 2934
1984----- 1983
2935 GACCCCTGCTCGTGTGTTACAGCGGGGTTTTTCTTGTGACAAGAAATCCTCACAAATACCGC 2994
1984----- 1983
2995 AGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGATCTCCGCTGTGCTTG 3054
1984----- 1983
3055 GCCAAAATTCGCAGTCCCCAACTCCCAATCACTCACCAACCTCTGCTCCAATTTGTC 3114
1984----- 1983
3115 CTGGTTATCGCTGATGTGTCTGCGGCTTTTATCATATATCTCTTTCATCTGCTGCTAT 3174
1984----- 1983
3175 GCCTCATCTTATTTGGTTCTTCTGGATTTATCAAGTATGTTGCCGCTTTGCTCTCTAA 3234
1984----- 1983
3235 TTCCAGGATCAACAACAACAGTACGGGACCATGCAAAACCTGACAGACTCCTGCTCAAG 3294
1984----- 1983
3295 GCAACTCTATGTTTCCCTCATGTTGCTGTACAAACCTACGGATGGAATTTGACCTGTA 3354
1984----- 1983
3355 TTCCCATCCATGCTCCTGGCTTTTCGAAATACCTATGGGAGTGGGCTCAGTCCGTT 3414
1984----- 1983
3415 TCTCTTGGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTTCGTAAGGCTTTTCCCCACATG 3474
1984----- 1983
3475 TTTGGCTTTTACGCTATATGATGATGTGTATTTGGGGGCAAGTCTGTACAGCATCGTGA 3534
1984----- 1983
3535 GTCCCTTTATACCGCTGTTACCAATTTTCTTTGCTCTGGGTATACATTTTAAAGATTCA 3594
1984-----CTAAGAAATTC 1994
3595 GACTCGAGCAAGTCTAGAAAGCGGCCCAAGATATCAAGATCCACTACGCGTTAGAGCT 3654
1995 GACTCGAGCAAGTCTAGAAAGCGGCCCAAGATATCAAGATCCACTACGCGTTAGAGCT 2054
3655 CGTGTATGAGCTCGACTGTGCTTCTAGTGTGACGCACTGTGTTGTTGCTCCCTCCCC 3714
2055 CGTGTATGAGCTCGACTGTGCTTCTAGTGTGACGCACTGTGTTGTTGCTCCCTCCCC 2114
3715 GTGCTTCTTGAACCTGGAAGTGGCACTCCCACTGTCTTCTTCTTAATAAATGAGGAA 3774
2115 GTGCTTCTTGAACCTGGAAGTGGCACTCCCACTGTCTTCTTCTTAATAAATGAGGAA 2174

QY 3775 ATTGCATCGCATTTGTTCTGAGTAGGTGTCATTCTATTCTCGGGGGTGGGGTGGGCAAGGAC 3834
DB |||||
QY 2175 ATTGCATCGCATTTGTTCTGAGTAGGTGTCATTCTATTCTCGGGGGTGGGGTGGGCAAGGAC 2234
DB |||||
QY 3835 AGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGAGCTCTTTCCGCTTCCTC 3894
DB |||||
QY 2235 AGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGAGCTCTTTCCGCTTCCTC 2294
DB |||||
QY 3895 GCTCACTGACTCGCTCGCTCGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA 3954
DB |||||
QY 2295 GCTCACTGACTCGCTCGCTCGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAA 2354
DB |||||
QY 3955 GCGGTAATACGGTTATCCACGAATCAGGGTAACGCGAAGAAACAATGTGACAAA 4014
DB |||||
QY 2355 GCGGTAATACGGTTATCCACGAATCAGGGTAACGCGAAGAAACAATGTGACAAA 2414
DB |||||
QY 4015 AGGCCAGCAAAAGGCCAGGAACCGTAAAGGCGGTTGCTGGCGTTTTTCCATAGGCT 4074
DB |||||
QY 2415 AGGCCAGCAAAAGGCCAGGAACCGTAAAGGCGGTTGCTGGCGTTTTTCCATAGGCT 2474
DB |||||
QY 4075 CCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGAC 4134
DB |||||
QY 2475 CCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGAC 2534
DB |||||
QY 4135 AGGACTATAAAGATACAGCGGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCC 4194
DB |||||
QY 2535 AGGACTATAAAGATACAGCGGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTCC 2594
DB |||||
QY 4195 GACCTTGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGGTTGCGCTTTC 4254
DB |||||
QY 2595 GACCTTGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGGTTGCGCTTTC 2654
DB |||||
QY 4255 TCAATGCTCACGCTGATAGTATCTCAGTTTCGGTGTAGGTGCTTCCAGTCCAAAGTGGGCTG 4314
DB |||||
QY 2655 TCAATGCTCACGCTGATAGTATCTCAGTTTCGGTGTAGGTGCTTCCAGTCCAAAGTGGGCTG 2714
DB |||||
QY 4315 TGTGCAGAACCCCGTTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGA 4374
DB |||||
QY 2715 TGTGCAGAACCCCGTTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGA 2774
DB |||||
QY 4375 GTCACACCCCGTTAAGACACAGCTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGATTAG 4434
DB |||||
QY 2775 GTCACACCCCGTTAAGACACAGCTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGATTAG 2834
DB |||||
QY 4435 CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGCGCTAACTACGGCTA 4494
DB |||||
QY 2835 CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGCGCTAACTACGGCTA 2894
DB |||||
QY 4495 CACTAGNAGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGGGAAG 4554
DB |||||
QY 2895 CACTAGNAGACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGGGAAG 2954
DB |||||
QY 4555 AGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTTG 4614
DB |||||
QY 2955 AGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTTG 3014
DB |||||
QY 4615 CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCCTTTGATCTTTTCTAC 4674
DB |||||
QY 3015 CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCCTTTGATCTTTTCTAC 3074
DB |||||
QY 4675 GGGGCTGAGCGCTCAGTGGNACGAAACTCACGCTTAAGGGATTTTGGTCATGAGATTATC 4734
DB |||||
QY 3075 GGGGCTGAGCGCTCAGTGGNACGAAACTCACGCTTAAGGGATTTTGGTCATGAGATTATC 3134
DB |||||
QY 4735 AAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAGTTTTTAAATCAATCTAAAG 4794
DB |||||
QY 3135 AAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAGTTTTTAAATCAATCTAAAG 3194
DB |||||
QY 4795 TATATATGATAACTTGTGCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTC 4854
DB |||||
QY 3195 TATATATGATAACTTGTGCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTC 3254
DB |||||

QY 4855 AGCGATCTGTCTATTCTTCTGTTTCATCTGATCTGCTGACTCCCGTCTGTTAGATAACTAC 4914
DB |||||
QY 3255 AGCGATCTGTCTATTCTTCTGTTTCATCTGATCTGCTGACTCCCGTCTGTTAGATAACTAC 3314
DB |||||
QY 4915 GATACGGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGCGAGACCCAGCTC 4974
DB |||||
QY 3315 GATACGGGAGGCTTACCATCTGCGCCAGTCTGCAATGATACCGCGAGACCCAGCTC 3374
DB |||||
QY 4975 ACGGCTCAGATTATCAGCAATAAACCCAGCCAGCGGAAAGGGCGAGGCGAAGTGG 5034
DB |||||
QY 3375 ACGGCTCAGATTATCAGCAATAAACCCAGCCAGCGGAAAGGGCGAGGCGAAGTGG 3434
DB |||||
QY 5035 TCTCGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAAG 5094
DB |||||
QY 3435 TCTCGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGAAGCTAGAGTAAG 3494
DB |||||
QY 5095 TAGTTGCGCAGTTAATAGTTTTCGCAACGTTGTTGCTTACAGGCAATCGTGTGTC 5154
DB |||||
QY 3495 TAGTTGCGCAGTTAATAGTTTTCGCAACGTTGTTGCTTACAGGCAATCGTGTGTC 3554
DB |||||
QY 5155 AGCTCGTCTGTTGTTGTTATGCTTCAGTCCCGTTCCCAACGATCAAGGGAGTTAC 5214
DB |||||
QY 3555 AGCTCGTCTGTTGTTGTTATGCTTCAGTCCCGTTCCCAACGATCAAGGGAGTTAC 3614
DB |||||
QY 5215 ATGATCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTTCGGTCTCCGATCGTGTCTAG 5274
DB |||||
QY 3615 ATGATCCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTTCGGTCTCCGATCGTGTCTAG 3674
DB |||||
QY 5275 AAGTAAGTTGGCGCAGTGTTCATCTCATGTTATGTCAGCACTGTCATAATCTCTTAC 5334
DB |||||
QY 3675 AAGTAAGTTGGCGCAGTGTTCATCTCATGTTATGTCAGCACTGTCATAATCTCTTAC 3734
DB |||||
QY 5335 TGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTG 5394
DB |||||
QY 3735 TGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTG 3794
DB |||||
QY 5395 AGAATAGTGTATGTCGGCGACCGAGTTGCTTCTTTCGCCGCGTCAATACGGGATAAATCCGC 5454
DB |||||
QY 3795 AGAATAGTGTATGTCGGCGACCGAGTTGCTTCTTTCGCCGCGTCAATACGGGATAAATCCGC 3854
DB |||||
QY 5455 GCCACATACGAGAACTTTTAAAAAGTGTCTCATTTGAAAAAGTCTTTCGGGGCGAAAACT 5514
DB |||||
QY 3855 GCCACATACGAGAACTTTTAAAAAGTGTCTCATTTGAAAAAGTCTTTCGGGGCGAAAACT 3914
DB |||||
QY 5515 CTCAGGATCTTACCGCTGTTGAGATCCAGTTTTCGATTAACCCACTCGTGCACCCACTG 5574
DB |||||
QY 3915 CTCAGGATCTTACCGCTGTTGAGATCCAGTTTTCGATTAACCCACTCGTGCACCCACTG 3974
DB |||||
QY 5575 ATCTTCAGCATCTTTTATCTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA 5634
DB |||||
QY 3975 ATCTTCAGCATCTTTTATCTTTCACAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAA 4034
DB |||||
QY 5635 TGCCGCAAAAAAGGGAATAAGCGGACACGGAAATGTTGAATACTCATACTCTTCTCTTTT 5694
DB |||||
QY 4035 TGCCGCAAAAAAGGGAATAAGCGGACACGGAAATGTTGAATACTCATACTCTTCTCTTTT 4094
DB |||||
QY 5695 TCAATATTATTGAAGCAATTTATCAGGGTTATTTGTTCTCATGCGGATACATATTTGAAATG 5754
DB |||||
QY 4095 TCAATATTATTGAAGCAATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATTTGAAATG 4154
DB |||||
QY 5755 TATTTAGAAAAATAACAAATAGGGGTTCCGGGCACATTTTCCCGGAAAAAGTGCACCTGA 5814
DB |||||
QY 4155 TATTTAGAAAAATAACAAATAGGGGTTCCGGGCACATTTTCCCGGAAAAAGTGCACCTGA 4214
DB |||||
QY 5815 CGTCTAAGAAAAACATTTATCATGACATTAACCTTATAAAAAATAGCGGTATCACGAGGCC 5874
DB |||||
QY 4215 CGTCTAAGAAAAACATTTATCATGACATTAACCTTATAAAAAATAGCGGTATCACGAGGCC 4274
DB |||||
QY 5875 CTTTCTGTC 5882
DB |||||
QY 4275 CTTTCTGTC 4282
DB |||||

RESULT 8

ACCT71561
ID ACC71561 standard; DNA; 6236 BP.
XX
AC ACCT71561;
XX
DT 10-JUL-2003 (first entry)
XX
XX VRC6802 (pVR1012x/s Lassa delta TW/h (codon optimised)) plasmid.
DE
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
XX W02003028632-A2.
PN
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030251.
PF
XX 01-OCT-2001; 2001US-0326476P.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI
XX WPI; 2003-371961/35.
DR
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 211-213; 219pp; English.
PS
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX
SQ Sequence 6236 BP; 1574 A; 1658 C; 1559 G; 1445 T; 0 U; 0 Other;

Query Match 41.1%; Score 2419.2; DB 10; Length 6236;
Best Local Similarity 70.9%; Pred. No. 1.4e-250;
Matches 3491; Conservative 0; Mismatches 1298; Indels 133; Gaps 16;

QY 1 TCGCGGGTTTCGGTGATGACGGTGAAAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
DB 1 TCGCGGGTTTCGGTGATGACGGTGAAAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60

QY 61 CAGCTTGCTGTGAAGCGGATGCGGGAGACAGACGCCCGTCAGGGCGGTCAAGCGGGTG 120
DB 61 CAGCTTGCTGTGAAGCGGATGCGGGAGACAGACGCCCGTCAGGGCGGTCAAGCGGGTG 120

QY 121 TTGGCGGGTTCGGGGCTGGCTTAACTATCGGGCATCAGACGAGATTGACTGAGAGTGC 180
DB 121 TTGGCGGGTTCGGGGCTGGCTTAACTATCGGGCATCAGACGAGATTGACTGAGAGTGC 180

QY 181 ACCATATGAAGCTTTTGGCAAAAGCCTAGGCGCTCCAAAAGCCCTCTCACTACTCTGG 240
DB 181 ACCATATGCGG-----TGTGAATACCGCACA 207

QY 241 AATAGCTCAGAGCGGAGCGGCTCGGCTCTGTCATAATAAAAAAATTAGTCAGCCA 300
DB 208 GATGCGTAAGGAGAAATATACCGCATCAG----- 235
QY 301 TGGGCGGAGAAATGGCGGAACTGGGCGGGAGGGAATTATTGGCTATTGGCCATTGTCAT 360
DB 236 -----ATTGGCTATTGGCCATTGTCAT 256
QY 361 AGCTTGCTATCATCATATATGTACATTTATTTGGCTCATGTCCCAATATGACCGCCA 420
DB 257 ACCTTGATTCATATATATATGTACATTTATTTGGCTCATGTCCCAATATGACCGCCA 316
QY 421 TGTTCAGATTGATTATTGACTAGTTAATAGTAATCAATTACGGGGTCAATTAGTTTCAT 480
DB 317 TGTTCAGATTGATTATTGACTAGTTAATAGTAATCAATTACGGGGTCAATTAGTTTCAT 376
QY 481 AGCCCATATATGAGTTCCGCGTTACATACTTACGGTAAATGGCGCGCTGGCTGACCG 540
DB 377 AGCCCATATATGAGTTCCGCGTTACATACTTACGGTAAATGGCGCGCTGGCTGACCG 436
QY 541 CCCAAGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAAGCCCAATA 600
DB 437 CCCAAGACCCCGCCCATTTGACGTCAATATGACGTATGTTCCCATAGTAAGCCCAATA 496
QY 601 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 660
DB 497 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 556
QY 661 CATCAAGTGTATCATATATGCAAGTCCGCGCTTATTTGACGTCAATGACGGTAAATGGCCC 720
DB 557 CATCAAGTGTATCATATGCAAGTCCGCGCTTATTTGACGTCAATGACGGTAAATGGCCC 616
QY 721 GCTGGCATTTATGCGCCAGTACATGACCTTACGGGACTTTTCCCTACTTTGGCAGTACATCTAC 780
DB 617 GCTGGCATTTATGCGCCAGTACATGACCTTATGGGACTTTTCCCTACTTTGGCAGTACATCTAC 676
QY 781 GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACACCAATAGGCGGTGA 840
DB 677 GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACATCAATAGGCGGTGA 736
QY 841 TAGCGGTTTGACTCAGCGGGATTTTCAAGTCTCCACCCCATTTGACGTCAATAGGAGTTTG 900
DB 737 TAGCGGTTTGACTCAGCGGGATTTTCAAGTCTCCACCCCATTTGACGTCAATAGGAGTTTG 796
QY 901 TTTTGGCACCAAAATCAACGGGACTTTCCAAATGTCGTAATTAACCCCGCCCGTTGACG 960
DB 797 TTTTGGCACCAAAATCAACGGGACTTTCCAAATGTCGTAATTAACCCCGCCCGTTGACG 856
QY 961 CAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTGTAGTGAAC 1020
DB 857 CAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTGTAGTGAAC 916
QY 1021 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGAC 1080
DB 917 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGAC 976
QY 1081 CGATCCAGCTCCGCGCGGGNAACGGTGCATTTGGAACCGGGATTTCCCGTGCACAGAT 1140
DB 977 CGATCCAGCTCCGCGCGGGNAACGGTGCATTTGGAACCGGGATTTCCCGTGCACAGAT 1036
QY 1141 GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCTATGCTATA 1200
DB 1037 GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCTATGCTATA 1096
QY 1201 CTGTTTTTGGCTTTGGGCGCTATACACCCCGC--TCCTTATGCTATAGGTGATGGTATAGC 1259
DB 1097 CTGTTTTTGGCTTTGGGCGCTATACACCCCGCCTTCTTATGCTATAGGTGATGGTATAGC 1156
QY 1260 TTAGCCCTATAGGTGTTGGTATTGACCATTTATGACCATCTCCCTTATTTGGTACGATCT 1319
DB 1157 TTAGCCCTATAGGTGTTGGTATTGACCATTTATGACCATCTCCCTTATTTGGTACGATCT 1216

QY 1320 TTCCATTACTAATCCATAACATGGCTCTTTGCCACAACCTATCTCTATTGGCTATATGCCA 1379
Db 1217 TTCCATTACTAATCCATAACATGGCTCTTTGCCACAACCTATCTCTATTGGCTATATGCCA 1276
QY 1380 ATACTCTGTCTTCAGAGACTGACCGGACTCTGTATTTTACAGGATGGGT-CCATTTT 1438
Db 1277 ATACTCTGTCTTCAGAGACTGACCGGACTCTGTATTTTACAGGATGGGTGCCATTT 1336
QY 1439 ATTATTACAAATTCACATATACAAACGCGCTCCCGTCCCGCAGCTTTTATTAA 1498
Db 1337 ATTATTACAAATTCACATATACAAACGCGCTCCCGCAGCTTTTATTAA 1396
QY 1499 CATAGCGTGGGATCTCC---GACATCTCGGTTAGCTGTTCCGACATGGCTCTTCTCG 1555
Db 1397 CATAGCGTGGGATCTCCACGCGAATCTCGGGTACGTGTTCCGACATGGCTCTTCTCG 1456
QY 1556 GTAGCGCGGAGCTTCCACATCCGAGCCCTGTGTCCTATCCGTCAGCGCTCATGTGCG 1615
Db 1457 GTAGCGCGGAGCTTCCACATCCGAGCCCTGTGTCCTATCCGTCAGCGCTCATGTGCG 1516
QY 1616 TCGGCACTCTTGTCTCTAACAGTGGAGCCAGACTTAAGGCACAGCAACAATGCCACCA 1675
Db 1517 TCGGCACTCTTGTCTCTAACAGTGGAGCCAGACTTAAGGCACAGCAACAATGCCACCA 1576
QY 1676 CCACCACTGTCCGCAACAGCCGCTGGCGGTAGGGTATGTGCTGAAATGAGCTCGGAG 1735
Db 1577 CCACCACTGTCCGCAACAGCCGCTGGCGGTAGGGTATGTGCTGAAATGAGCTGGAG 1636
QY 1736 ATTGGCTCGCAC-CTGGAGCGAGATGGAAGACTTAAGGCAGCGGCGAGGAAGATGCAG 1794
Db 1637 ATTGGCTCGCACCGGTGACCGAGATGGAAGACTTAAGGCAGCGGCGAGGAAGATGCAG 1696
QY 1795 GCAGCTGAGTGTGTATTCTGATAAGAGTCAGAGTAACTCCGTTGCGGTGCTGTAA 1854
Db 1697 GCAGCTGAGTGTGTATTCTGATAAGAGTCAGAGTAACTCCGTTGCGGTGCTGTAA 1756
QY 1855 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTGTTGCTGCGCGCGCGGCCACAGACATA 1914
Db 1757 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTGTTGCTGCGCGCGCGGCCACAGACATA 1816
QY 1915 ATAGCTGACAGACTTCTCTTTTCATAGGTCTTTTCTGAGTCACCGTCGTC 1974
Db 1817 ATAGCTGACAGACTTCTCTTTTCATAGGTCTTTTCTGAGTCACCGTCGTC 1876
QY 1975 GACGAATTCAGCAATCATGATG-CAATGAAGAGAGGCTCTGCTGTGTGCTGCTG 2033
Db 1877 GACACGTGTGATCAGATATCGCGCGCTCTAGAGATATCGCGCCATGGGCCAGATCGT 1936
QY 2034 TGTGGAGCAGTCTTCTGTTTCGCCAGGCTAGCGAAACCCACG-TCACCGGGGAGTGC 2092
Db 1937 GACCTTTCTCCAGGAGTGCCTATGTGATCGAGGAGGTGATGAACATCGTGTGATCGC 1996
QY 2093 CGGCCACACTGTCTGTGATTGTAGCTCTCTCGCACAGGCGCCAAAGCAGAACGTCCA 2152
Db 1997 CTTGAGGCTGTGGCGTGTGAAGGCTGTACAACTTCGCCACCTGCGGCCCTGTGGG 2056
QY 2153 GCTGATCAACACCAACCGCAGTTGGCACTCAATAGCAGCGCCCTGAACCTGCAATGATG 2212
Db 2057 CTTGGTGACCTTCCTGTCTGTGCGCAGGAGCTGCACACCCAGCTGTACAAAGGCGT 2116
QY 2213 CTTCAACACCGGCTGTGGCAGGCTTTTCTATACCAAGATTCACTCTTCAGGCTG 2272
Db 2117 GTACGAGCTGCAGACCTTGAGCTGAACATGGAGACCTTGAAACATGACCAATGCCCTGAG 2176
QY 2273 TCTGAGAGGCTAGCAGCTGCGACCCCTTACCGATTTTGGACAGGCTGGGGCCCTAT 2332
Db 2177 CTGCAACAAGAACACACGACCACTACATCATGTGTGGGCAACGACCGCTGAGCT 2236
QY 2333 CAGTTATGCCAAGCGGCGCCCGACCGAGCCCTACTGCTGGCACTTACCCCCCAA 2392
Db 2237 AACCTGACCAACACGAGCATCATCAACCAAGATTCTGCAACCTGAGCGAGCCCAAA 2296
QY 2393 ACCTTGGGTATTGTGCCCCG---GAAGAGTGTGTGCTCGGTATATTGCTTCACTCC 2449

Db 2297 GAAGAACCTGTATCAGACACGCGCTGATGAGCATCATCAGACCTTCCACTGAGCATCCC 2356
QY 2450 CAGCCCCGTGTGTGGGAAACGACACAGGTGCGGCGCGCCACCTACAGCTGGGGTGA 2509
Db 2357 CAACTTCAACCACTACGAGGCCATGAGCTGCGACTTCAACGGGGCAAGATCAGCGTGA 2416
QY 2510 AAATGATACGAGCGTCTTGTCTTAACTAACTAATACAGGCCACCGCTGGGCAATTTGTTGG 2569
Db 2417 GTACAACCTGAGCCACAGCTAGCCGCGACCGCCAAACCACTGCGGACCGTGGCCAA 2476
QY 2570 TTGTACTCTGATGAACCTCAACTGGATTCAAAAGTGTGCGAGCGCTCTTGTGTCTAT 2629
Db 2477 CGCGTGTGACAGCTTCTCATGAGGATGCGCTGGGGCGGAGCTACATCGCCTGAGACAG 2536
QY 2630 CGAGGGGGCGGCAACAAACACCTGTCACCTGCCCCCACTGATTTGCTCCGCAAGCATCCGA 2689
Db 2537 CGGAGGGGCAACTGGGACTGTCATATGACAGCTACCACTACCTGATCATCCAGAACAC 2596
QY 2690 CGGCACTACTCTCGGTGGGCTCCGCTCCCTGATGATCACCAGAGTGGCTGGTCTGACTA 2749
Db 2597 CACTGGGAGGACCACTGCGCAGTTGAGCAGGCGCCAGCCCCATCGGCTACTTGGGCCCTGCT 2656
QY 2750 CCGGTATAGCTTTGSCATTATCTTGTATACCATCAACTACACCATATTTAAATCAGGAT 2809
Db 2657 GAGCCAGAGGACCAAGGACATCTCATCAGCAGAGGCTGCTGGGCACCTTCACTTGGAC 2716
QY 2810 GTACGTGGGAGGGGTGCAACACAGGCTGGAAGCTGCTGCAACTGGACGCGGGCGAAGC 2869
Db 2717 CTTGAGCGACGAGGCGCAAGACACACCCGCGGCTACTGCTGACAGGTGGATGCT 2776
QY 2870 TTGCGATCTGGAAGATAGGAGCAGGTCCGAGATCGATATGATGGAACATCACATCAGATT 2929
Db 2777 GATCGAGGCGGAGCTGAAGTGTCTGGCAACACCCGCGTGGCCAACTGCAACGAGAAGCA 2836
QY 2930 CTTAGGACCCCTGCTCGTGTATACGCGGGTCTTTCTTTGTTGCAAGAATCTCTCAAT 2989
Db 2837 CGACGAGGAGTCTTGGCAGCATGCTGAGGCTGTTCGACTTCAACAGCAGGCCATCCAGAG 2896
QY 2990 ACCGAGAGCTTAGACTCGTGGTGGACTTCTCTCAATTTTCTAG-GGGGATCTCCGCTGT 3048
Db 2897 GCTGAAGGCGGAGGCCCCAGATGAGCATCCAGCTGATCAACAAAGCCGTGAACGCCCTGAT 2956
QY 3049 GTCTTGGCCAAATTTGCGAGTCCCAACCTCAATCACTCAACCACTCTGCTCTCCAA 3108
Db 2957 CAACGACCACTGATCATGAAGAACCACTGAGGAGCATCATGGGCATCCCTACTGCAA 3016
QY 3109 TTTGTCTCTGTTATCGCTGGATGTCTCGCGC-----GTTTTATCATATTCTCTTCTAT 3163
Db 3017 CTACAGCAAGTACTGTTACTGAAACCAACACCCAGCGGAGGACAGCTGCCAAGTG 3076
QY 3164 CTTGCTGTATGCTCTCATCTTCTATTG-GTTCTTCTGATTTAAGGTATGTTGCCCCG 3222
Db 3077 CTGGCTGGTGAGCAACGCGAGCTTACCTGAAACAGAGACCCACTTCAGCGACGACATCGAGCA 3136
QY 3223 TTTGTCTCTAATTCAGGATCAACAAACAGTAGGAGCAACCTGCAAACTCTGACGA 3282
Db 3137 GCAGGCGCAACATGATCAGGATCAGGATGCTGAGAGAGGAGTACATGGAGGAGGCGCAA 3196
QY 3283 CTCCTGCTCAAGGCAACTCTATTGTTCC---CTCATGTTGCTGTACAAACCTTACGGATG 3339
Db 3197 GACCTGAACAGTGGGATCCAGATCTGCTGTGCTTCTAGTTGCCAGCATCTGTTGTTT 3256
QY 3340 GAAATGCACTGTATTTCCTATCCCATCCCATCGTCTGGGCTTTTCGCAAAATACCTATGGAGT 3399
Db 3257 GCCCTCCCGGTCCTCTTGACCTGGAAGTGCCTCCACTCCCACTGCTCTTTCTTAAT 3316
QY 3400 GGGCTCAGTCGTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTAGTGGTTCGTAG 3459
Db 3317 AAATGAGGAAATTTGCATCGCATTTGCTGAGTAGGTGTCTATTCTATTTGGGGGGTGGG 3376
QY 3460 GGCTTCCCGCACTGTTTGGCTTTTCAGCTATATGATGATGTGTTATTTGGGGGCCAAGTC 3519

CC	determine the efficacy of the S gene segment of Seoul virus to provide
CC	protective immunity the gene segment was subcloned into the naked DNA
CC	expression vector pWRG7077 downstream of the cytomegalovirus immediate
CC	early promoter to yield pWRG/SEO-S. The construct was then used to
CC	challenge hamsters. Postchallenge serum samples from pWRG/SEO-S
CC	vaccinated hamsters were positive for glycoprotein and nucleocapsid
CC	antibodies
XX	
SQ	Sequence 6050 BP; 1634 A; 1416 C; 1454 G; 1546 T; 0 U; 0 Other;
	Query Match 40.8%; Score 2399.4; DB 3; Length 6050;
	Best Local Similarity 72.1%; Pred. No. 1.8e-248;
	Matches 3457; Conservative 0; Mismatches 1096; Indels 240; Gaps 16;
QY	337 ATATTGGCTATTGGCCATTGACATCGTTGTATCTATATCATATAATATGACATTTATATT 396
DB	1263 AATATTGGCTATTGGCCATTGGCATTGCTGTATCTATATCATATAATATGACATTTATATT 1322
QY	397 GGCTCATGTCCAAATATAGCCGCCATGTGACATTGATTTAGTACTAGTTATTAATAGTAA 456
DB	1323 GGCTCATGTCCAAATATAGCCGCCATGTGACATTGATTTAGTACTAGTTATTAATAGTAA 1382
QY	457 TCAATTACGGGGTCATTAGTTTCATAGCCCATATATGAGTTCCGCGTTACATAACTTACG 516
DB	1383 TCAATTACGGGGTCATTAGTTTCATAGCCCATATATGAGTTCCGCGTTACATAACTTACG 1442
QY	517 GTAAATTGGCCGCTGGCTGACCGCCCAAGCACCCCCCGCCCATTTGACGTCATTAATGACG 576
DB	1443 GTAAATTGGCCGCTCG-TGACGCCCAAGCACCCCCCGCCCATTTGACGTCATTAATGACG 1501
QY	577 TATGTTCCCATAGTAAGCCCAATAGGACTTTCATTTGACGTCATTAAGGTTGGAGTATTTA 636
DB	1502 TATGTTCCCATAGTAAGCCCAATAGGACTTTCATTTGACGTCATTAAGGTTGGAGTATTTA 1561
QY	637 CGGTAAATGCCCCACTTGGCAGTACATCAAGTGTATCATATGCAAGTCC-GCCCCCTAT 695
DB	1562 CGGTAAATGCCCCACTTGGCAGTACATCAAGTGTATCATATGCAAGTCCGCGCCCCCTAT 1621
QY	696 TGAAGTCAATGACGGTAAATGGCCCCGCTGGCATTAATGCCAGTACATGACCTTACGGGA 755
DB	1622 TGAAGTCAATGACGGTAAATGGCCCCGCTGGCATTAATGCCAGTACATGACCTTACGGGA 1681
QY	756 CTTTCTCTACTTGGCAGTACATCTACGTTATTAGTTCATCGCTTATACCATGGTGTGCGGTT 815
DB	1682 CTTTCTCTACTTGGCAGTACATCTACGTTATTAGTTCATCGCTTATACCATGGTGTGCGGTT 1741
QY	816 TTGGCAGTACACCAATGGGGCGTGGATAGCGGTTTGGACTCACGGGGATTTCCAAAGTCTCCA 875
DB	1742 TTGGCAGTACACCAATGGGGCGTGGATAGCGGTTTGGACTCACGGGGATTTCCAAAGTCTCCA 1801
QY	876 CCCCATTGACGTCATATGGGAGTTTGTATTGGCACCAAAATCAACGGGACTTTCCAAAATG 935
DB	1802 CCCCATTGACGTCATATGGGAGTTTGTATTGGCACCAAAATCAACGGGACTTTCCAAAATG 1861
QY	936 TCGTAATAACCCGCCCGCTTACGCAAAATGGGGGTAGCGGTGACGGTGGAGGTCTA 995
DB	1862 TCGTAATAACCCGCCCGCTTACGCAAAATGGGGGTAGCGGTGACGGTGGAGGTCTA 1921
QY	996 TATAAGCAGAGCTCGTTTATGTGAACCGTCAGATCGCTGGAGACGCCATCCACGCTGTTT 1055
DB	1922 TATAAGCAGAGCTCGTTTATGTGAACCGTCAGATCGCTGGAGACGCCATCCACGCTGTTT 1981
QY	1056 TGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCGCGCCGGGAACGGTGCATTTGG 1115
DB	1982 TGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCGCGCCGGGAACGGTGCATTTGG 2041
QY	1116 AACCGGATTTCCCGTCCCAAGAGTGACGTAAAGTACCGCTATAGACTCTATAGGACAC 1175
DB	2042 AACCGGATTTCCCGTCCCAAGAGTGACGTAAAGTACCGCTATAGACTCTATAGGACAC 2101
QY	1176 CCCTTTGGCTCTTATGATGCTACTACTGTTTGGCTTGGGGCTATACACCCCGCTCC 1235
DB	2102 CCCTTTGGCTCTTATGATGCTACTACTGTTTGGCTTGGGGCTATACACCCCGCTCC 2161

QY	1236 TTATGCTATAGTGTATGGTATAGCTTAGCTATAGTGTGGGTATTGACCATTTATGAC 1295
DB	2162 TTATGCTATAGTGTATGGTATAGCTTAGCTATAGTGTGGGTATTGACCATTTATGAC 2221
QY	1296 CACTCCCTATTGGTGACCATATCTTCCATTAATCAATCAATCAATCAATCAATCAATCA 1355
DB	2222 CACTCCCTATTGGTGACCATATCTTCCATTAATCAATCAATCAATCAATCAATCAATCA 2281
QY	1356 ACTATCTCTATTGGCTATATGCAATACCTCTCTCTCAGAGACTGACACGGACTCTGTA 1415
DB	2282 ACTATCTCTATTGGCTATATGCAATACCTCTCTCTCAGAGACTGACACGGACTCTGTA 2341
QY	1416 TTTTTCACAGATGGGGT-CCATTTATTTTACAAATTCACATATACAAACAGCCGCTCC 1474
DB	2342 TTTTTCACAGATGGGGTCCCATTTATTTTACAAATTCACATATACAAACAGCCGCTCC 2401
QY	1475 CCGGTGCCCGCAGTTTATTAATAACATACGCTGGGATCTCC--GACATCTCGGGTACGT 1531
DB	2402 CCGGTGCCCGCAGTTTATTAATAACATACGCTGGGATCTCCACGGGAATCTCGGGTACGT 2461
QY	1532 GTTCCGGACATGGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCTCC 1591
DB	2462 GTTCCGGACATGGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCTCC 2521
QY	1592 ATCCGTCGAGCGGCTCATGGTCTGCTGGCAGCTCTTGTCTCTTAACAGTGGAGGCCAGAC 1651
DB	2522 ATCCGTCGAGCGGCTCATGGTCTGCTGGCAGCTCTTGTCTCTTAACAGTGGAGGCCAGAC 2581
QY	1652 TTAGGCACAGCAATGCCACACACACAGTGTGCCGACAAAGGCCGTGGCGGTAGGGT 1711
DB	2582 TTAGGCACAGCAATGCCACACACAGTGTGCCGACAAAGGCCGTGGCGGTAGGGT 2641
QY	1712 ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCAAGATGGAAGACTTAA 1771
DB	2642 ATGTGTCTGAAATGAGCTCGGAGATTGGGCTCGCACCTGGACGCAAGATGGAAGACTTAA 2701
QY	1772 GGCAGCGGACAGAAAGATGACGGCAGCTGAGTTGTTGTATTCTGATAGAGTCAAGGT 1831
DB	2702 GGCAGCGGACAGAAAGATGACGGCAGCTGAGTTGTTGTATTCTGATAGAGTCAAGGT 2761
QY	1832 AACTCCGCTTGGGCTGCTGTTAAACGCTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGC 1891
DB	2762 AACTCCGCTTGGGCTGCTGTTAAACGCTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGC 2821
QY	1892 TCCGCGCGCGCCACACAGACATAATAGCTGACAGACTAAACAGACTGTTCTTTCCATGGG 1951
DB	2822 TCCGCGCGCGCCACACAGACATAATAGCTGACAGACTAAACAGACTGTTCTTTCCATGGG 2881
QY	1952 TCTTTTCTGAGTCACTCCGTCGTCGAGCAATTCAAGCAATCATGGATGCAATGAGAGAGG 2011
DB	2882 TCTTTTCTGAGTCACTCCGTCGTCGAGCAATTCAAGCAATCATGGATGCAATGAGAGAGG 2937
QY	2012 GCTCTGCTGTGCTGCTGCTGTGTGGAGCAGTCT--TCGTTTCCGCCACGCGCTAGCGA 2068
DB	2938 ACTCCCTAAGAGCTACTACTACTAAACAGAAATGCGAACTATGGAAGAAATCCAGAGA 2997
QY	2069 AACCCACGTCACCGGGGGAAGTCCCGGCCACACTGTGTCTGGATTGTTAGCCTCTCTCGC 2128
DB	2998 GAAATCAGTGTCTCACGAGGGGAGCTTGTGTATAGCAGCGCAGAGGTCAAGGATGCAGAA 3057
QY	2129 ACCAGGCGCAAGCAGACAGCTCCAGCTGATCAACACCAACGGCAGTTGGCACTCTCAATAG 2188
DB	3058 AAGCAGTATGAGAGGATCTTGATGACTTTAAACAGAGGGGCACTGCATGATCGGGAGGT 3117
QY	2189 CACGGCCCTGAACTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCA 2248
DB	3118 GTGCGAGCTTCAATACAAATCAAAATTTGATGATTTGAAGGCCCACTTGCACAGATTG 3177
QY	2249 CCACAAAGTTCACTCTTTCAGGGTGTCTCTGAGAGGTAGCCAGCTGCCGACCCCTTACCGA 2308
DB	3178 CAGCAGGGAAGAACATCCGGGGAGGACCGGATCTCTACAGGGGTAGAGCCAGGTGATCAT 3237

DT	10-JUL-2003 (first entry)
XX	VRC6500 (pVR1012-VP35) plasmid.
DE	Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW	viral infection; filovirus; circular; cyclic; ds.
XX	Synthetic.
OS	WO2003028632-A2.
XX	10-APR-2003.
PN	24-SEP-2002; 2002WO-US030251.
PD	01-OCT-2001; 2001US-0326476P.
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI	WPI; 2003-371961/35.
XX	New bimodal priming and boosting compositions, useful as viral vaccines,
DR	specifically for eliciting an immune response against a filovirus or a
XX	disease caused by infection with filovirus.
PS	Claim 1; Page 169-170; 219pp; English.
XX	The present invention relates to a bimodal priming composition and
CC	boosting composition for priming and boosting an immune response to an
CC	antigen in an individual. The compositions comprise (a) a priming
CC	composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC	encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC	virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC	DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC	composition comprised of a replication-deficient adenovirus, comprising a
CC	nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC	paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC	epitope-bearing domain, or a replication deficient adenovirus selected
CC	from constructs ACC71521-ACC71563. The compositions are useful as viral
CC	vaccines, specifically for eliciting an immune response against a
CC	filovirus or a disease caused by infection with filovirus e.g. Ebola
CC	virus infection. This plasmid has an insert of VP35 from Ebola cloned
CC	from pGEM 3zf(+)/VP35 (#1213) and is derived from a pVR1012 backbone
XX	Sequence 6148 BP; 1636 A; 1513 C; 1448 G; 1551 T; 0 U; 0 Other;
SQ	
Query Match	40.7%; Score 2395.6; DB 10; Length 6148;
Best Local Similarity	71.6%; Pred. No. 4.6e-248;
Matches 3516; Conservative	0; Mismatches 1194; Indels 204; Gaps 20;
QY	1 TCGCCGGTTTCGGTGATGACGGTGAAGAACCTCTGCACATCGAGCTCCCGGAGACGGTCA 60
DB	
DB	1 TCGCCGGTTTCGGTGATGACGGTGAAGAACCTCTGCACATCGAGCTCCCGGAGACGGTCA 60
QY	61 CAGCTTGCTGTAAAGCGGATCCGGGAGCAGACAGCCGTCAGGGCGCTCAGCGGGTG 120
DB	
DB	61 CAGCTTGCTGTAAAGCGGATCCGGGAGCAGACAGCCGTCAGGGCGCTCAGCGGGTG 120
QY	121 TTGGCGGGTGTCCGGGCTGGCTTAACATATCGGGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB	
DB	121 TTGGCGGGTGTCCGGGCTGGCTTAACATATCGGGCATCAGAGCAGATTGTACTGAGAGTGC 180
QY	181 ACATATGAAGCTTTTTCGAAAGCGCTAGGCCCTCCAAAAGCGCTCTCACTACTCTTGG 240
DB	
DB	181 ACCATAATGCGG-----TGTGAATAACCGCAC 207
QY	241 AATAGCTCAGAGCCGAGCGCCCTCGGCCCTCTGCATTAATAAAAAAATTAGTCACCCA 300
DB	
DB	208 GATGCGTAAGGAGAAATACCGCATCAG----- 235

|||||
1277 ATACTCTGCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGGTCCCATTT 1336
QY
1439 ATTATTTTACAAATTCACATATACAAAGCCGCTCCCGTCCCGCAGTCTTTTATTTAA 1498
Db
1337 ATTATTTACAAATTCACATATACAAAGCCGCTCCCGTCCCGCAGTCTTTTATTTAA 1396
QY
1499 CATAGCGTGGGATCTCC---GACATCTCGGGTACGTGTTCCGGACATGGGCTCTTCCG 1555
Db
1397 CATAGCGTGGGATCTCCAGCGAATCTCGGTACGTGTTCCGGACATGGGCTCTTCCG 1456
QY
1556 GTAGCGCGGAGCTCCACATCCGAGCCCTGCTGCCATCCGTCACGGGCTCATGTCG 1615
Db
1457 GTAGCGCGGAGCTCCACATCCGAGCCCTGCTGCCATCCGTCACGGGCTCATGTCG 1516
QY
1616 TCGCAGCTCTCTTCTCTAAAGTGGAGGCCAGACTTAAAGCACACACAATGCCACCA 1675
Db
1517 TCGCAGCTCTCTTCTCTAAAGTGGAGGCCAGACTTAAAGCACACACAATGCCACCA 1576
QY
1676 CCACAGTGTCCGCAACAAGCCGCTGGGTAGGGTATGTGCTGAAATGAGCTCGGAG 1735
Db
1577 CCACAGTGTCCGCAACAAGCCGCTGGGTAGGGTATGTGCTGAAATGAGCTCGGAG 1636
QY
1736 ATTGGGCTCGCAC-CTGGAGCGAGATGGAAGACTTAAAGCAGCGGCAGAGAGATGCAG 1794
Db
1637 ATTGGGCTCGCACAGCGCTGACGAGATGGAAGACTTAAAGCAGCGGCAGAGAGATGCAG 1696
QY
1795 GCAGCTGAGTGTGTTATTTCTGATAAGAGTCAGAGGTAATCTCCGTTGCGGTGCTGTAA 1854
Db
1697 GCAGCTGAGTGTGTTATTTCTGATAAGAGTCAGAGGTAATCTCCGTTGCGGTGCTGTAA 1756
QY
1855 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGGCCACACAGACATA 1914
Db
1757 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGGCCACACAGACATA 1816
QY
1915 ATAGCTGACAGACTTAACAGACTGTTCTTCCATGGGTCTTTTCTCAGTCAACGTCGTC 1974
Db
1817 ATAGCTGACAGACTTAACAGACTGTTCTTCCATGGGTCTTTTCTCAGTCAACGTCGAA 1876
QY
1975 GAC---GAATTCACAGCAATCATGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTG 2030
Db
1877 TTCTCTAGCACTCGAAGCTTATGTCTTCAATGTAAAGAAAGCTGGTCTAACAGATG 1936
QY
2031 CTGTGTGAGCAGTCTTCTGTTTCGCCAGCGTAGCGAAACCCACGTC-ACCGGGGGAAG 2089
Db
1937 ACAACTAGAACAAAGGCGAGGGGCCATACTGCGGCCACGACTCAAAACGACAGAAATGCCA 1996
QY
2090 TGC CGGCCACACTGTCTGATTTGTTAGCTCTCTCGCACAGGCGCCAAAGCAGAACGT 2149
Db
1997 GGCCCTGAGCTTTGGGCTGGATCTCTGAGCAGCTAATGACCGGAAGAAATTCCTGTAAGC 2056
QY
2150 CCAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCACGCGCCCTGAACTGCAATGA 2209
Db
2057 GACATCTTCTGTATATTGAGAACAAATCCAGGATTTATGCTAGCATCCCAAAATGCAACAA 2116
QY
2210 TAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATACCAACAAGTTCAATCTTTCAGG 2269
Db
2117 ACGAAGCCAAACCGAAGACGCGCA---ACAGTCAAAACCCAAACGACCCCAATTTGCAAT 2173
QY
2270 CTGTCTGAGAGCTAGCAGCTGCCGACCCCTTACCGATTTTACAGCGGCTGGGGCCG 2329
Db
2174 CATAGTTTGGAG----AGGTAGTACAACAATATGGCTTCATTGGCTTACTGTTGTGCAAC 2228
QY
2330 TATCAGTTATGCCAAGGAGCGGCCCGACACGAGCGCCCTACTGTGGCAGCTACCCCCC 2389
Db
2229 AACAAACCATCGCATCAGAAATCATAGAACACGCAATACGAGCTTTGAGAATGGTCTAA 2288
QY
2390 AAAACCTTGGGGTATTGTGCCCGGAGAGAGTGTGTGTGTCGGTATATTTGCTTCACTCC 2449
Db
2289 AGCCAGTTTATGATATGGCAAAACAAATCTCTCTCATTTGAACAGGGTTTGTGCTGAGATGG 2348
QY
2450 CAGCCCGTGGTGGTGGGAACGACAGAGGTGGGGCGGCCACCTACAGCTGGGGTGA 2509
|||||

Db
2349 TTGCAAAATATGATCTTCTGGTGATGACAAACGGTCCGGCAACAGCAACCGCTGCGGCAA 2408
QY
2510 AAATGATACGGACGCTTTCGTCCTTTAAACAATACAGGCCACCGCTGGGCAAATGGTTCGG 2569
Db
2409 CTGAGG-----CTTATTTGGGCCGAACATGTGTCAACACCACTTGGACCATCAGCTTTAT 2461
QY
2570 TTGTACTCTGGATGAATCAACTGGATTACAAAGTGTGCGGAGCGCTTCTTGTGTCTCAT 2629
Db
2462 GAAGAAGTGCAGTTTCGGGTAAGATTGAATCTAGAGATGAGACCGTCCC---TCAAAG 2517
QY
2630 CCGAGGGCGGGCAACAACACCTGCACTGCCCACTGATTTGCTTCCGCAAGCATCCGGA 2689
Db
2518 TGTTAGGAGGCAATCAACAAATCTAAACAGTACCACT----- 2554
QY
2690 CCGCACAATCTCTCGTTCGGTCCCGTCCCTGGATACACCCAGGTGCTGCTGCACTA 2749
Db
2555 --TCACTAATCAGGAAAAATTTTGGAAACCTGCAATTTCCGCAAGGATTTGAGAAAA 2612
QY
2750 CCGGTATAGGCTTTGGCATTATCCTTTGTACCATCACTACACATATTTAAATCAGGAT 2809
Db
2613 TTATGTATGATCACTTGCCTGTTTGGAACTGCTTTTCCACCAATTTAGTACAGTGAAT 2872
QY
2810 GTACGTGGAGGGTCTGAACACAGGCTGGAAGCTGCTGCAACTGGACCGGGCGAAACG 2869
Db
2673 GTAAATTTGGAAAAAGATAGCAACTCATTTGGACATCATCTCATGTGAGTTCCAGGCCAGCC 2732
QY
2870 TTGCGATCTGGAAGATAGGACAGGTCGAGATCGATATGGAACATCACATCAGGAT 2929
Db
2733 TGGCTGAAGGAGACTCTCTCAATGTGCCTAATTCAAATTTACAAAGAGTTCCAATCT 2792
QY
2930 CCTAGACCCCTGCTGTTTACAGCGGGGTTTCTTGTGACAGAAATCCTCAAT 2989
Db
2793 TCCAAGATGCTGCTCATCTGTCACTCACATCCGCTTTCAGGTGACATTTCCCGAGCTT 2852
QY
2990 ACCGACAGCTTAGACTCTGCTGGTGA-CTTCTCTCAATTTTCTAGGGGATCTCCCGTGT 3048
Db
2853 GCCAGAAAGCTTGGTCCAGTCCACCATCGCCCAAGATTTGATCAGGTTGGGATGTT 2912
QY
3049 GTCTTGGCCAAATTCG---CAGTCCCAACCTCCAATCACTCACCAACCTCTCTGCTC 3105
Db
2913 TTTTCAGCTTCAAGATGGTAAACACTTGGACTCAAAATTTGAGCCCAATCTCCCTTCCCT 2972
QY
3106 CAATTTGCTCTGTTATCGCTGATGCTGCTGGGGTTTATCATATTTCTCTTCTTCACTC 3165
Db
2973 CCGAAAGAGCGGAATAATAGCAGAGCTTCAACTGCTGAACTATAGGGTACGTTACAT-- 3030
QY
3166 TGCTGTATGCTCTCATCTTCTTATTTGTTTCTTCTGATTTATCAAGGTATGTTGCGCGT 3225
Db
3031 -----TAAATGATACACTTGTGATATCAGCCCTGGATAATATAAGTCAATTAACGACC 3084
QY
3226 GTCTCTAATTTCCAGATCAACAAACAGTACGGGACATGCAAAAACCTTGACCACTC 3285
Db
3085 AAGATAAAATTTGTCATATCTCGTAGCAGCTTAAATATAAATGTAATAGGAGCTATAT 3144
QY
3286 CTGCTCAAGCAGCACTCATATGTTTCCCTCATGTTGCTGTAACAAAACCTACGGATGGAAT 3345
Db
3145 CTCTGACAGGGGATCCAGATCTGCTGTGCTCTTCTAGTTGCCAGC----- 3189
QY
3346 GCACCTGTATTTCCCATCCCATCGTCTGGGCTTTTCGCAAAATACCTATGAGGAGTGGGCT 3405
Db
3190 ---CATCTGTTTGGCCCTCCCGCTGCTTCTTTCAGCCCTGGAAGGTGCCACTCCCA 3246
QY
3406 CAGTCTGTTTCT-CTTGGCTCAGTTTACATAGTGCATTTGTTTTCAGTGGTTCGTAGGGCT 3464
Db
3247 CTGTCTTTTCTTAATAAATGAGAAATTCGATTCGATTTGTCTGAGTAGGTGTCTTCTA 3306
QY
3465 TCCCCCAGCTGTTTGGCTTTTTCAGCTATGATGATGTTGTTTGGGGCCAAAGTCTGTAC 3524
Db
3307 TTTCTGGGGGGTGGGGTGGGGGACAGCAGAGGGGGAGGATTTGGGAAGACAATAGCAGGC 3366
QY
3525 AGCATCGTGAAGTCCCTTTTATACCGCTGTTTACCAATTTCTTTTGTCTCTGGGTATACAT 3584
Db
3367 ATGCTGGGATGCGGTGGGCTCTATGGGTACCCAGGTGCT-----GAAGAATTG 3415

Matches 3498; Conservative 0; Mismatches 1280; Indels 177; Gaps 19;

QY	1	TCGCGGCTTTCGGTGATGACGGTGAAACCTCTGACACATGACAGTCCCGGAGACGGTCA	60
Db	1	TCGCGGCTTTCGGTGATGACGGTGAAACCTCTGACACATGACAGTCCCGGAGACGGTCA	60
QY	61	CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGGCTCAGCGGCTG	120
Db	61	CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGGCTCAGCGGCTG	120
QY	121	TTGGCGGCTGTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGC	180
Db	121	TTGGCGGCTGTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAGAGTGC	180
QY	181	ACCATATGAAGCTTTTGGCAAAAGCCTAGGCGCTCCAAAAAGCCCTCCTCACTACTTCTGG	240
Db	181	ACCATATGGG-----TGTGAATACCGGCACA	207
QY	241	AATAGCTCAGAGCCGAGGCGGCTCGGCTCTGCTATTAATAAAAAAATTAGTCAGCCA	300
Db	208	GATGGTAAGGAGAAATACCGCATCAG-----	235
QY	301	TGGGCGGAGAAATGGGCGGAACTGGGCGGAGGGAATTATGGCTATTGGCCATTGCAAT	360
Db	236	-----ATTGGCTATTGGCCATTGCAAT	256
QY	361	ACCTTGATCTATATCATATATATGATACATTTATATTGGCTCATGTCCAATATGACGCCA	420
Db	257	ACGTTGTATCCATATCATATATATGATACATTTATATTGGCTCATGTCCAATATGACGCCA	316
QY	421	TGTTGACATTGATTATGACTAGTATTAAATAGTAAATCAATTACGGGTCATTAGTTTCAT	480
Db	317	TGTTGACATTGATTATGACTAGTATTAAATAGTAAATCAATTACGGGTCATTAGTTTCAT	376
QY	481	AGCCCATATATGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCGCTGCTGACCG	540
Db	377	AGCCCATATATGAGTTCGCGCTTACATAACTTACGGTAAATGGCCCGCTGCTGACCG	436
QY	541	CCCAACGACCCCGCCCATGAGCTCAATATGACGTATGTTCCCATAGTAAAGCCCAATA	600
Db	437	CCCAACGACCCCGCCCATGAGCTCAATATGACGTATGTTCCCATAGTAAAGCCCAATA	496
QY	601	GGGACTTTCATGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACATTTGGCAGTA	660
Db	497	GGGACTTTCATGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACATTTGGCAGTA	556
QY	661	CATCAAGTGTATCATATGCCAAGTCCGCCCTTATTGACGTCAATGACGGTAAATGGCCC	720
Db	557	CATCAAGTGTATCATATGCCAAGTACGCCCTTATTGACGTCAATGACGGTAAATGGCCC	616
QY	721	GCCTGGCATATGCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC	780
Db	617	GCCTGGCATATGCCAGTACATGACCTTATGGGACTTTCCCTACTTTGGCAGTACATCTAC	676
QY	781	GTATTAGTCAATCCCTATTACATGGTGTGCGGTTTTGGCAGTACACCAATGGCGGTGA	840
Db	677	GTATTAGTCAATCCCTATTACATGGTGTGCGGTTTTGGCAGTACATCAATGGCGGTGA	736
QY	841	TAGCGGTTTGACTCAGCGGGAATTCCAAAGTCTCCACCCCATGACGTCAATGGGAGTTTG	900
Db	737	TAGCGGTTTGACTCAGCGGGAATTCCAAAGTCTCCACCCCATGACGTCAATGGGAGTTTG	796
QY	901	TTTTGGGACCAAAATCAACGGGACTTTCCAAATATGCTGAATAAACCCCGCCCGCTTGACG	960
Db	797	TTTTGGGACCAAAATCAACGGGACTTTCCAAATATGCTGAATAAACCCCGCCCGCTTGACG	856
QY	961	CAAAATGGGAGTGGGCTGACCGTGGAGTCTATATAGCAGAGCTCGTTTTAGTGAAC	1020
Db	857	CAAAATGGGAGTGGGCTGACCGTGGAGTCTATATAGCAGAGCTCGTTTTAGTGAAC	916
QY	1021	CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGAC	1080
Db	917	CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGAC	976

QY	1081	CGATCCAGCCTCCGCGGCCGGAAACGGTCATTGGAAACGGGATTTCCCGTGCCAGAGT	1140
Db	977	CGATCCAGCCTCCGCGGCCGGAAACGGTCATTGGAAACGGGATTTCCCGTGCCAGAGT	1036
QY	1141	GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGATGCTATA	1200
Db	1037	GACGTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGATGCTATA	1096
QY	1201	CTGTTTTTGGCTTGGGCGCTATACACCCCGCG-TCCTTATGCTATATAGGTGATGTTATAG	1259
Db	1097	CTGTTTTTGGCTTGGGCGCTATACACCCCGCTTCTTATGCTATATAGGTGATGTTATAG	1156
QY	1260	TTAGCCTATAGGTGCGGTTATTGACCATTATGACCACCTCCCTTATTTGGTGACGATCT	1319
Db	1157	TTAGCCTATAGGTGCGGTTATTGACCATTATGACCACCTCCCTTATTTGGTGACGATCT	1216
QY	1320	TTCCATTACTAATCCATAACATGGGCTTTGGCAACAATCTCTATTGGCTATATGCCA	1379
Db	1217	TTCCATTACTAATCCATAACATGGGCTTTGGCAACAATCTCTATTGGCTATATGCCA	1276
QY	1380	ATACTCTGCTTTCAGAGACTGACACGGACTCTGTATTTTTACAGGATGGGT-CCATTT	1438
Db	1277	ATACTCTGCTTTCAGAGACTGACACGGACTCTGTATTTTTACAGGATGGGTCCCATT	1336
QY	1439	ATTATTTACAAATTCACATATACACACCGCGCTCCCGCTGCCCGCAGTTTTATTAA	1498
Db	1337	ATTATTTACAAATTCACATATACACACCGCGCTCCCGCTGCCCGCAGTTTTATTAA	1396
QY	1499	CATAGCTGGGATCTCC--GACATCTCGGGTACGTGTTCCGGACATGGGCTCTTCTCG	1555
Db	1397	CATAGCTGGGATCTCCAGCGGAATCTCGGTAAGTGTTCGGACATGGGCTCTTCTCG	1456
QY	1556	GTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCCATCCGTCACGCGCTCATGTGCG	1615
Db	1457	GTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCCATCCGTCACGCGCTCATGTGCG	1516
QY	1616	TGGGAGCTCTTGTCTCTTAAAGTGGGCGCAGACTTAGGCACAGCAAAATGCCACCA	1675
Db	1517	TGGGAGCTCTTGTCTCTTAAAGTGGGCGCAGACTTAGGCACAGCAAAATGCCACCA	1576
QY	1676	CCACAGTGTGCGCACAAAGCGCGTGGCGGTAGGTTATGTCTGAAAATAGACTCGAG	1735
Db	1577	CCACAGTGTGCGCACAAAGCGCGTGGCGGTAGGTTATGTCTGAAAATAGACTCGAG	1636
QY	1736	ATTGGGCTCGCAC-CTGGACCGCAGATGGAAGACTTAAAGCAGCGGCAGAAAGATGCG	1794
Db	1637	ATTGGGCTCGCACCGCTGACGAGATGGAAGACTTAAAGCAGCGGCAGAAAGATGCG	1696
QY	1795	GCAGCTGAGTTGTTGTTATTTCTGATAGAGTCAGAGTAACTCCGTTGCGGTGCTGTTAA	1854
Db	1697	GCAGCTGAGTTGTTGTTATTTCTGATAGAGTCAGAGTAACTCCGTTGCGGTGCTGTTAA	1756
QY	1855	CGGTGGAGGCGAGTGTAGTCTGACAGTACTCGTTGCTCGCGCGCGCCACACAGACATA	1914
Db	1757	CGGTGGAGGCGAGTGTAGTCTGACAGTACTCGTTGCTCGCGCGCGCCACACAGACATA	1816
QY	1915	ATAGCTGACAGACTAACAGACTGTTCTTTTCAATGGGTTCTTTCTGCACTACCGTCGTC	1974
Db	1817	ATAGCTGACAGACTAACAGACTGTTCTTTTCAATGGGTTCTTTCTGCACTACCGTCGTC	1876
QY	1975	GACGAATTCAGCAATCATGGATGC---AATGAGAGAGGGCTCTGCTGTGTGCTGCTG	2030
Db	1877	GACAGTGTGATCAGATATCGCGCGCGCTCTAGAAATTTTAGGATTTGGCTTTTAGAATTC	1936
QY	2031	CTGTGTGGAGCAGTCTTCTGTTTCCGCCAGCGCTAGGAAACCCACGCTACCCGGGGAGT	2090
Db	1937	ACTACTAGTTAGGAATTCCTAAATCATGGGCGAGATTATACATTTCTTTCAAGAGTGCC	1996
QY	2091	GCGGCGCACATGTGTCTGGATTGTTAGC-CTCCTTCGACCAAGGCGCCCAAGACGAT	2149
Db	1997	ACATGTAATAGGAGAGTCAAGCAATTTGTCTAATTTGGCTTTCTCTATTGGCAATCTTT	2056

QY 2150 CCAGCTGATCAACACCAACGCGAGTTGGCACTCAATAGACAGCGCCCTGAACTGCAATGA 2209
DB 2057 GAAGGGCTGTATAAATCGCTACATGTGGGATTTATGGATTTGTTGCCCTTTTATTTCTT 2116
QY 2210 TAGGCTCAACACCGGCTGGTTGGCAGGCTTTTCTATCACCAACAAGTTCAACTCTTCAGG 2269
DB 2117 GTGTGGCAAGCTGTGTTCCTTAACCCCTTAAAGGGGATATGAGCTGCAAACTTAGAATT 2176
QY 2270 CTGTCTGTAGAGGCTAGCAGCTGCCAGCCCTTACCGATTTTGACAGGGCTGGGCCC 2329
DB 2177 AATATGAGAGCCCTAAACATGACCATGCCCTTATCATGACCAAGAACAGAGTCAATCA 2236
QY 2330 T---ATCAGTTATGCCAAACGGAAGCGGCCCGACAGCGC---CCCTACTGTGCGACTA 2383
DB 2237 TTACATAAGAGTGGGCAATGAGACTGTGATTTAGAATTGACTTTAACTAACACACGACATTAT 2296
QY 2384 CCCCCAAAACCTTGGGTATTTGGCCGGAAGAGTGTGTGGTCCGGTATATTGCTT 2443
DB 2297 AATCAACAAATTTTGCACATTTATCCGATGCTCAAAAAGAAATCTTTATGATCATGCTCT 2356
QY 2444 CACTCCAGCCCGGTGTGTGGGAACGACCGACAGGTGCGGCGGCCACCTACAGCTG 2503
DB 2357 CATGAGCTCATCTCAACATTCATCTATCTCATCTTCCAACTTCAATGATGAGCCAT 2416
QY 2504 GGGTGAATATGATACGAGCTCTTGTCTTAAATAATACAGGCCACCGCTGGGCAATTG 2563
DB 2417 GAGTTGTGATTTCAATGGAGGGAATCAGTGTGCAATAACAACCTCTCTCATTTCTTATGC 2476
QY 2564 GTTCGGTTGATCTGAGTGAATCAACTGATTAACCAAGATGCGGAGCGCTCTCTTG 2623
DB 2477 TGGGGATGCGGCCGCAACACTGTGGGACAGTTGCGCAACGAGAGTGTGCAACATTTATGAG 2536
QY 2624 TGTCACTCGGAGGGCGGCAACAAACACCTGTGCACTGCCACCTGATTTGCTTCGCAAGCA 2683
DB 2537 AATGGCTTGGGGTGGAGATACATTTGATTTAGACTCAGGAAGGAACTTGGGACTGTAT 2596
QY 2684 TCCGGACGCAATATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTGCCCTGGT 2743
DB 2597 AATGACAGCTACAGTACCTGATTAATTTCAAAATACAAATGAGGAGGACCACTGCCAATT 2656
QY 2744 CGACTACCGTATAGGCTTTGG--CATATCTTTGATCACTCAACTACACCATATTAA 2801
DB 2657 CTCGAAGCCGTCTCTATCGGTACCTTGGCCCTTTTGTCAAAAGGACAAAGATATATA 2716
QY 2802 ATCAGATGTACGTGGAGGGTTCGAACACAGGTGGAAGCTGCCCTGCACTGACGCGG 2861
DB 2717 TATAAGTAGAGGCTCTTGGGGACCTTTCAC-----CTGGAATGTCAGATTCTG 2766
QY 2862 GGGCAACGTTTGGATCTGGAAGATAGGGAACAGGTTCGAGATCGATATGGAAGAACATCACA 2921
DB 2767 AGGCAATGAACACACAGGTGTTATTTGTTTAAACAGGTGATGCTAATTTGAAGCAGAAC 2826
QY 2922 TCAGGATTCCTAGAACCCCTGCTGCTGTTACAGCGGGGTTTTTCTGTTGACAAAGATC 2981
DB 2827 TCAAGTGTTTTGGGAATACAGCTGTGGCAAAATGCAATGAGAAGCATGATGAGGAGTTT 2886
QY 2982 CTCACAATACCGCAGAGCTAGACTGTGTGTGACTTCTCTCAATTTTCTAGGGGATCT 3041
DB 2887 GTGACATGCTGAGATTGTTTGAATTTCAACGAAGCAATCCGTAGTTGAAGCTGAGG 2946
QY 3042 CCGGTGTCTTGGCCAAAATTCGAGTCCGCCAACCTCCAACTCACTCAACCAACCTCTGT 3101
DB 2947 CCCAGATGATTTCAATTAATAATAAGCCGTGAATGCTTAATCAATGATCAATTAA 3006
QY 3102 CTCCA-----ATTGTCTGTGTTATTCGCTGAGATGTGCTGGGGCTTTAT 3148
DB 3007 TCATGAAGAACCAATTTAAGAGACATCATGGGCAATTCCTACTGCAATTAACGAAGTAT 3066
QY 3149 CATATCTCTTCTATCTGCTGATGCTCACTTCTTATTTGTTCTTCTGATATCA 3208
DB 3067 GGTACCTTAATCATACTAGTAGGGGAGAACTCACTACCAAGTGTGTGGCTTATATCCA 3126
QY 3209 AGGTATGTTGCCCGTTTGTCTCTAATTTCCAGGATCAACAAACACGAGTACGGGACCATG 3268

DB 3127 ATGGGTCAATCTTAATGAAGCCAGTCTCTGATGACATAGAACAGCAAGCCGACAATA 3186
QY 3269 CAAAACCTGCACGACTCTG-----CTCAAGGCACTCTATGTGTTTCCCTCATGTT 3318
DB 3187 TGATCACAGAGATGCTTTCAGAAAGAAATACATTTGAAAGACAAGGGAACACGCCCCTTGAGG 3246
QY 3319 GCTGTACAAAACCTACGGATGGAATTCG---ACCTGTATTCCTCATCCCATCGTCTCTGGG 3375
DB 3247 GATCCAGATCTGCTGTGCTCTTAGTTGGCAGCCATCTGTGTTTGGCCCTTCCCGCTGC 3306
QY 3376 CTTTCCAAAATACCTATGAGGAGTGGCTCAGTCCGTTTCT--CTTGGCTCAGTTTACTA 3434
DB 3307 CTTCTTGAACCCCTGGAAGGTGCACTCCCACTGCTCTTCTTAATAAAATGAGGAAATTG 3366
QY 3435 GTGCCATTTGTTTCAAGTGTTCGTAGGCTTCCCCCACTGTTTGGCTTTTCAGCTATATGG 3494
DB 3367 CATCGCATTTGCTGAGTAGGTGTCATTCTTGGGGGTGGGGTGGGCGAGGACAGCA 3426
QY 3495 ATGATGTGTATTTGGGGGCAAGTCTGTACAGCATCGTGAAGTCCCTTTTATACCGCTGTTA 3554
DB 3427 AGGGGAGGATTTGGGAAGACAATAGCAGGATGCTGGGATGCGGTGGGCTCTATGGGTA 3486
QY 3555 CCAATTTTCTTTTG---TCTCTGGGTATACATTTAAAGAAATTCAGACTCGAGCAAGTCTAG 3611
DB 3487 CCAGGTGCTGAAGAAATTGACCCGTTCTCTGGGCCAGAAAGAAAGCAGGACATCCCC 3546
QY 3612 AAAGGCGGCCAAGATATCAAGGATCCACTACGCTTAGAGTCTGCTGATCAGCCTCGAC 3671
DB 3547 TTCTCTGACACACCCCTGTCCACGCCCTGTGTTCTTAGTTCAGGCCCACTCATAGGAC 3606
QY 3672 -----TGTGCTTCTAGTTGCGAGCCATCTGTGTTGTCCTCCCTCCCGCTGCC 3719
DB 3607 ACTCATAGCTCAGAGGGCTCCGCTTCAATCCACCCGCTAAAGTACTTTGGAGCGGTCT 3666
QY 3720 TTCTTTGACCCCTGGAAGTGGCACTCCCACTGCTCTTCTTCTTAATAAAATGAGGAAATTCG 3779
DB 3667 CTCCCTCCCTCATCAGCCCAACCAAACTAGCTCCAGAGTGGGAAGAAATTTAA 3726
QY 3780 ATCGCATTTCTGAGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGCGAGGACAGCA 3839
DB 3727 GCAAGATAGGCTATTAAAGTGCAGAGGAGAGAAAATGCTCCCAACATGTGAGGAAGTAT 3786
QY 3840 GGGGAGGATTTGGGAAGACAATAGCAGGATGCTGGGAGCTTCCGCTTCTCTCGCTCA 3899
DB 3787 GAGAGAAATCATAAGATTTTAAAGCCATCATG--GCTTAACTTTCGCTTCTCTCGCTCA 3844
QY 3900 CTGACTCGCTCGCTCGTTCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGG 3959
DB 3845 CTGACTCGCTCGCTCGTTCGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGG 3904
QY 3960 TAATACGTTTTCACAGAAATCAGGGATTAACGAGGAAAGAAATGCTGAGCAAAAGGCC 4019
DB 3905 TAATACGTTTTCACAGAAATCAGGGATTAACGAGGAAAGAAATGCTGAGCAAAAGGCC 3964
QY 4020 AGCAAAAGCCAGGAACCGTAAAGGCGCGTGTGCTGCGGTTTTTCCATAGGCTCCGCC 4079
DB 3965 AGCAAAAGCCAGGAACCGTAAAGGCGCGTGTGCTGCGGTTTTTCCATAGGCTCCGCC 4024
QY 4080 CCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTCAAGTGGCGGCAAAACCCGACAGGAC 4139
DB 4025 CCGCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTCAAGTGGCGGCAAAACCCGACAGGAC 4084
QY 4140 TATAAGATACAGGGGTTTTCCCTGGGAAGTCCCTCGCTGCGCTCTCTGTTTCCGACCC 4199
DB 4085 TATAAGATACAGGGGTTTTCCCTGGGAAGTCCCTCGCTGCGCTCTCTGTTTCCGACCC 4144
QY 4200 TGGCGCTTACCGATACCTGTCCGCTTCTTCCCTTCCGGAAGCGTGGCGCTTCTCAAT 4259
DB 4145 TGGCGCTTACCGATACCTGTCCGCTTCTTCCCTTCCGGAAGCGTGGCGCTTCTCAT 4204
QY 4260 GCTCAGCTGTAGGTATCTCAGTTCCGTTAGTTCGTTCCCTCAAGCTGGGCTGTGTC 4319

Db 6025 ||||| AGGTGCTGCTCCAAAGCTGGCTGTGTGCAGCAACCCCGCTTACGCCGACCGTGGC 6084
Qy 4350 CTTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCCGCACTGG 4409
Db 6085 CTTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTAAGACACGACTTATCCGCACTGG 6144
Qy 4410 CAGCAGCACTGGTAAACAGAGTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCT 4469
Db 6145 CAGCAGCACTGGTAAACAGAGTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCT 6204
Qy 4470 TGAAGTGGTCCCTAACTACGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGC 4529
Db 6205 TGAAGTGGTCCCTAACTACGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTGC 6264
Qy 4530 TGAAGCCAGTCTTACCTTCGGAAGAGTGTGCTAGCTCTTCATCCGCAACCAACACCG 4589
Db 6265 TGAAGCCAGTCTTACCTTCGGAAGAGTGTGCTAGCTCTTCATCCGCAACCAACACCG 6324
Qy 4590 CTGGTAGCGGTGGTTTTTTTGTTCAGCAGCAGAGTATACGCGCAGAAAAAGGATCTC 4649
Db 6325 CTGGTAGCGGTGGTTTTTTTGTTCAGCAGCAGAGTATACGCGCAGAAAAAGGATCTC 6384
Qy 4650 AAGAAGATCTTTGATCTTTTACCGGGTCTGACGCTCAGTGGAGCAAACTCAGTT 4709
Db 6385 AAGAAGATCTTTGATCTTTTACCGGGTCTGACGCTCAGTGGAGCAAACTCAGTT 6444
Qy 4710 AAGGATTTTGGTTCATCAGATTAATCAAAAAGGATCTTCACCTAGATCTTTTAAATTA 4769
Db 6445 AAGGATTTTGGTTCATCAGATTAATCAAAAAGGATCTTCACCTAGATCTTTTAAATTA 6504
Qy 4770 AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 4829
Db 6505 AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 6564
Qy 4830 GCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTGTTGTTCAATCATAGTTGCCT 4889
Db 6565 GCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTGTTGTTCAATCATAGTTGCCT 6624
Qy 4890 GACTCCCCGTCGTGTAGATAACTACGATACGGAGGGCTTACCATTCTGGCCCCAGTGGCTG 4949
Db 6625 GACTCCCCGTCGTGTAGATAACTACGATACGGAGGGCTTACCATTCTGGCCCCAGTGGCTG 6684
Qy 4950 CAATGATACCGCAGAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCCAG 5009
Db 6685 CAATGATACCGCAGAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACCAGCCAG 6744
Qy 5010 CCGGAAGGCGCAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTA 5069
Db 6745 CCGGAAGGCGCAGCGCAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTA 6804
Qy 5070 ATTGTTGCCGGAAGCTAGAGTAGTATGCTGCGCAGTTAATAGTTTGGCAAGCTTTGTTG 5129
Db 6805 ATTGTTGCCGGAAGCTAGAGTAGTATGCTGCGCAGTTAATAGTTTGGCAAGCTTTGTTG 6864
Qy 5130 CCATTGCTACAGGCATCGTGGTGTACGCTCGCTGGTTGGTATGCTTCAATCAGCTCCG 5189
Db 6865 CCATTGCTACAGGCATCGTGGTGTACGCTCGCTGGTTGGTATGCTTCAATCAGCTCCG 6924
Qy 5190 GTTCCCAACGATCAAGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCT 5249
Db 6925 GTTCCCAACGATCAAGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCT 6984
Qy 5250 CTTCCGCTCCGATCGTTGTCAGAGTAAGTTGGCCGAGTGTATCACTCATGTTA 5309
Db 6985 CTTCCGCTCCGATCGTTGTCAGAGTAAGTTGGCCGAGTGTATCACTCATGTTA 7044
Qy 5310 TGGCAGCACTGCATAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTCTGTGACTG 5369
Db 7045 TGGCAGCACTGCATAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTCTGTGACTG 7104
Qy 5370 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTGTCTTTGCC 5429
|||||

Db 7105 GTGAGTACTCAACCAAGTCAATTTCTGAGAAATAGTGTATGCGCGCACCGAGTGTCTTTGCC 7164
Qy 5430 CGCGCTCAATACGGGATAATACCGGCCACATAGACAGAACTTTTAAAGTGTCTCATCATTG 5489
Db 7165 CGCGCTCAACGGGATAATACCGGCCACATAGACAGAACTTTTAAAGTGTCTCATCATTG 7224
Qy 5490 GAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGCA 5549
Db 7225 GAAAAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGCA 7284
Qy 5550 TGTAAACCACTCGTGCACCCAACTGATCTTACGATCTTTTACTTTTCAACGAGTTTCTG 5609
Db 7285 TGTAAACCACTCGTGCACCCAACTGATCTTACGATCTTTTACTTTTCAACGAGTTTCTG 7344
Qy 5610 GGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAAT 5669
Db 7345 GGTGAGCAAAAAACAGGAAGCAAAATGCCGCAAAAAAGGAATAAGGGCGACACGGAAT 7404
Qy 5670 GTTGAATACCTCATCTCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGTGTC 5729
Db 7405 GTTGAATACCTCATCTCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTATGTGTC 7464
Qy 5730 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGTTCCGCGCA 5789
Db 7465 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGTTCCGCGCA 7524
Qy 5790 CATTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACCATTTATTCATGACATTAACCT 5849
Db 7525 CATTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACCATTTATTCATGACATTAACCT 5884
Qy 5850 ATAAAAATAGGCGTATCACGAGCGCCCTTTTCGTC 5982
Db 7585 ATAAAAATAGGCGTATCACGAGCGCCCTTTTCGTC 7617
RESULT 15
ABK91498
ID ABK91498 standard; DNA; 6577 BP.
XX
AC ABK91498;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #50.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US025721.
XX
PR 14-AUG-2000; 2000US-0225097P.
PR 14-NOV-2000; 2000US-0252115P.
PR 28-MAR-2001; 2001US-0279257P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX
PI Nabel GJ, Huang Y;
XX
DR WPI; 2002-454594/48.
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunization for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection.
XX

Db 1757 CGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTGTGCTCGCGCGCGCCACAGACATA 1816
Qy 1915 ATAGCTGACAGACTAAACAGACTGTCTCTTTCCATGGGTCTTTTCTGCACTCAGTCGTCGTC 1974
Db 1817 ATAGCTGACAGACTAAACAGACTGTCTCTTTCCATGGGTCTTTTCTGCACTCAGTCGTC 1876
Qy 1975 GACGAATTCAGCA---ATCATGATGCAATGAAGAGGGCTCTGCTGTGCTGCTGTC 2031
Db 1877 GACACGTGTGATCAGATATCGCGCGCTCTAGACACCATGCGCTGAAGGAGATACC 1936
Qy 2032 TGTGTGGAGCAGTCTTCGTTTCCGCCAGCGCTAGCGNAACCCACGTCACCGGGGAGTG 2091
Db 1937 AGCACCTGTGGCGTGGGGCTGGCGTGGGGCCACATGCTGCTGGGATGCTGATGATCT 1996
Qy 2092 CCGGCCACACTGTGTCTGATTTTGTAGCTCTCGCACCGAGCGCCCAAGCAGAAAGTCC 2151
Db 1997 GCAGCGCCACGAGC-----TAGCAGCTTGAAGTGCACCGACCTGAAGACGACA 2047
Qy 2152 AGCTGATCAACACCAACGCGAGTTGGACCTCAATAGCACGGCCCTGAACCTGCAATGATA 2211
Db 2048 CCAACACCAACAGCAGCGCGCGCATGATCATGAGAGAGGGCGAGATCAAGAACTGCA 2107
Qy 2212 GCCTCAACACCGGCTGTTGGAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCT 2271
Db 2108 GCTTCAACATCAGCACAGCATCCGCGGCAAGGTGCAGAAAGGATACGCCCTTCTTACA 2167
Qy 2272 GTCTCAGAGGCTAGCAGCTGCGACCCCTTACCGATTTTGACAGGGCTGGGGCCCTA 2331
Db 2168 AGTGGACATCATCCCATCGACAAACGACACACAGCTACAGCTTACAGCTGCAACA 2227
Qy 2332 TCAGTTATGCAACGGAAGCGGCCCGACAGCGCCCTACTCTGCTGGCACTACCCCCCA - 2390
Db 2228 CCAGCGTATCACCCAGGCTGCCCAAGGTGAGCTTCGAGGCCATCCCCCAACCACTACT 2287
Qy 2391 --AAACCTTCGCTATTTGTCGCGGAGAGTGTGTGTCGCTGATTTATGCTTCACTC 2448
Db 2288 GCGCCCGCGCTTGGCATCTGAAAGTGCAGAGGCAAGAAAGTTCAAGCGCAAGGGCC 2347
Qy 2449 CCAGCGCCGCTG---GTGGAAACGACGACAGGTGCGGGCGGCCACCTACAGCTGG 2504
Db 2348 CTTGCAACCAAGTAGCAGCTGAGTGACACCGCATCCCGCGCGGTGGTGAACACC 2407
Qy 2505 GGTGAATAATGATACGAGCTTTCTGCTTAAACATAACAGGCCACCGCTGGCAATTGG 2564
Db 2408 AGCTGCTGTAACCGTAGCTGCGGAGGAGGTGGTGTATCCCGAGCGCTAACTTGG 2467
Qy 2565 TTGCTGTGCTGATGATCACTCACTGATTCACCAAGTGTGGAGGGCT--CCTT 2622
Db 2468 CCGACAAACGCAAGGTGATCATCTGCGAGCTGAACGAGAGCGTGGAGATCAACTGCACCC 2527
Qy 2623 GTGTGATCGAGGGGGGCAACACACCTGCACTGCCCACTGATTGCTTCCGCAAGC 2682
Db 2528 GCGCCAAACAAACACCGCAAGAGATTCACATCGGCCCGCGCGGCGCTTCTACACCA 2587
Qy 2683 ATCCGACGCGCACATCTCTCGGTGGGCTCCGGTCCCTGGATCAACCCAGGTGCTGG 2742
Db 2588 CCGCGAGATCATCGCGACATCCGCGAGGCCCACTGCAACCTGAGCGCGCAAGTGA 2647
Qy 2743 TCAGCTACCGGTATAGCTTTGGATATTCCTTGTACCATCAACTACACCATATTTAAA 2802
Db 2648 ACACACCTTGAACAAGATCTGTATCAAGTGCAGGAGTTCGGCAACAAGACCATCG 2707
Qy 2803 TCAGGATGTACGTGGAGGGGTGAAACACAGGCTGGAAGCTGCTGCAACTGGACGCGG 2862
Db 2708 TGTTCAGACACAGCGGGCGGACCCCGAGATCTGACCCACAGCTTCACTGCGGCG 2767
Qy 2863 GCGAAGTTGCGATCTGGAAGATAGGACAGGTCCGAGATCGATATGGAGAACATCACAT 2922
Db 2768 GCGAGTTCTTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGTTTCAACAGCACCT 2827
Qy 2923 CAGGATTCCTAGACCCCTGCTGCTGTTACAGCGGGGTTTTTCTGTTGACAAAGATCC 2982
Db 2828 GGAGCACCGAGGCGACAAAC-----ACCGAGGGCGAGCGACACCATCACCCCTGCCCT 2881

Qy 2983 TCACAATACGAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGATCTC 3042
Db 2882 GCGCATCAAGCAGATCATCAACATGTGCAAGAGGTGGCAAGCCATGTACGCCCCC 2941
Qy 3043 CGGTGTCTTGGCCAAAATTCGAGTCCCAACCTCCAACTCAATCACTCAACAACTCTGTGTC 3102
Db 2942 CCATCAGCGCCAGATCCGCTGAGCAGCAACATCACCGGCTGCTGCTCGCCGACG 3001
Qy 3103 CTCCAATTTGCTGCTGTTATCGCTGATGTGCTGCGGGTTTTTATCATATCTCTCTCA 3162
Db 3002 GCGCAACAGCAACAAAGAGCGAGATCTTCCGTCGCGCGCGGCGGACATGCGCA 3061
Qy 3163 TCTGCTGTATGCTCATCTTCTTATGCTTCTTCTGATATC---AAGTATGTTG 3218
Db 3062 ACTGCGCAGCGAGCTGTACAGTCAAGTGTGTAAGATCGAGCCCTGGCGTGGCCC 3121
Qy 3219 CCGTTTGTCTTAATTCAGGATCAACAAACAGTACGGGACATGCAAAACCTGC 3278
Db 3122 CCACCAAGGCCAAGCTTACCGTCCAGGCCCGCAGCTGCTGAGCGGATCGTCAGCAGC 3181
Qy 3279 ACGACTCTGCTCAAGGCAACTTATGTTTCCCTCATGTTGCTGTACAAAACCTACGAT 3338
Db 3182 AGAACAACTGTGCGCGCATCGAGGCCAGCAGCACCTGCTGCACTGACCGTGTGGG 3241
Qy 3339 GGAATTTGCACTGTATTCCTATCCCATCCCTGCTCTGG---GCTTTCGCAAAATACCTATGG 3395
Db 3242 GCATCAAGCAGCTGCGAGCCCGCACCTGCGCTGGAGCGCTACCTGAGGACCGACG 3301
Qy 3396 GAGTGGGCTCAGTCCGTTTCTTGGCTCAGTTTACTAGTGCATTTGTTCAAGTGTTC 3455
Db 3302 AGCTCAGCAGATCTGGAACCAACACCTGATGAGTGGGACCGCGAGATCAACAACT 3361
Qy 3456 GTAGGCTTTCCCACTGTTGGCTTTTCCAGCTATATGATGATGATGTTTGGGGCCA 3515
Db 3362 ACACAGCTGATTCACAGCTGATCGAGGAGAGCCAGAACACGACGAGAAAGCAGC 3421
Qy 3516 AGTCTGTACAGCATCGTGAAGTCCCTTTATACCGCTGTACCAATTTTCTTCTCTCTGG 3575
Db 3422 AGGAGCTGTGGAGCTGGACAAGTGGCGCAGCTGTGGAACCTGGTTCAACATCACCACT 3481
Qy 3576 GTATACATTTAAGAAATTCAGATCTGAGCAAGTCTAGAAAGCGCGCAAGATATCAAGGA 3635
Db 3482 GGTGTGTGTACATCAAGCTGTTTATCATGATCGTGGCGCGCTGTGGGCTCGCATCG 3541
Qy 3636 TCCACTACCG--TTAGAGCTCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCGAGCAT 3694
Db 3542 TGTTCGCCGTGTGAGCATCTGAGGATCCAGATCTGCTGTGCTTCTAGTTGCGAGCAT 3601
Qy 3695 CTGTTGTTTGGCCCTCCCGCTGCTTCTTGAACCTGGAAGGTGCCACTCCCACTGTCC 3754
Db 3602 CTGTTGTTTGGCCCTCCCGCTGCTTCTTGAACCTGGAAGGTGCCACTCCCACTGTCC 3661
Qy 3755 TTTCTTAATAAATGAGGAAATTCGATTCGATTTGTCTGAGTAGTGTCTATTTATCTGG 3814
Db 3662 TTTCTTAATAAATGAGGAAATTCGATTCGATTTGTCTGAGTAGTGTCTATTTATCTGG 3721
Qy 3815 GGGTGGGTGGGCGAGGACAGCAAGGGGAGATTTGGGAAGACAATAGCAGGATGCTG 3874
Db 3722 GGGTGGGTGGGCGAGGACAGCAAGGGGAGATTTGGGAAGACAATAGCAGGATGCTG 3781
Qy 3875 GGGAGC----- 3880
Db 3782 GGGATCGGTGGGCTCTATGGGTACCCAGGTGCTGAAGAATTGACCCGGTTCCTCTGGG 3841
Qy 3881 ----- 3880
Db 3842 CCAGAAAGAGCAGGACATCCCTTCTCTGTGACACACCTGTCCACGCCCTGTTCT 3901
Qy 3881 ----- 3880
Db 3902 TAGTTCACGCCCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCAATCCCAC 3961

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2005, 22:14:31 ; Search time 929 Seconds
(without alignments)
10360.150 Million cell updates/sec

Title: US-10-715-665-6

Perfect score: 5882

Sequence: 1 tcgcgcgtttcgtgatgac.....tatcacgagcccttcgtc 5882

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq*

2: /cgn2_6/prodata/1/ina/5B_COMB.seq*

3: /cgn2_6/prodata/1/ina/6A_COMB.seq*

4: /cgn2_6/prodata/1/ina/6B_COMB.seq*

5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq*

6: /cgn2_6/prodata/1/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5882	100.0	5882	4	US-09-721-480-6
2	4617.2	78.5	5459	4	US-09-721-480-4
3	4190	71.2	5128	4	US-09-721-480-2
4	2658.4	45.2	4276	4	US-09-721-480-1
5	2265.4	38.5	8575	5	PCT-US92-08258-6
6	2256.6	38.4	6565	3	US-09-299-141-1
7	2229.2	37.9	6280	3	US-08-893-327-19
8	2225.6	37.8	6280	3	US-08-893-327-17
9	2220.6	37.8	6253	3	US-08-893-327-15
10	2166.4	36.8	4328	3	US-09-132-808-1
11	2166.4	36.8	4328	3	US-08-910-647-2
12	2166.4	36.8	4328	3	US-08-920-925-2
13	2166.4	36.8	4328	3	US-09-620-260-1
14	2166.4	36.8	4328	4	US-09-620-259-1
15	2156.6	36.7	5618	3	US-08-799-569-1
16	2156.6	36.7	5618	3	US-09-570-546-1
17	2156.6	36.7	5618	4	US-09-146-072A-1
18	2147.6	36.5	5703	1	US-08-467-420A-50
19	2147.6	36.5	5703	1	US-08-470-110A-50
20	2147.6	36.5	5703	1	US-08-667-769A-50
21	2147.6	36.5	5703	2	US-08-940-371-50
22	2147.6	36.5	5703	3	US-08-637-647-50
23	2147.6	36.5	5703	5	PCT-US95-17082A-50
24	2141	36.4	6710	4	US-09-355-040-28
25	2109	35.9	9600	3	US-08-910-647-1
26	2109	35.9	9600	3	US-09-620-925-1
27	2065.4	35.1	5952	2	US-08-663-998-2

ALIGNMENTS

RESULT 1

US-09-721-480-6
; Sequence 6, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5882
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pCMV-II-E2661-sag
; NAME/KEY: CDS
; LOCATION: (1992)..(3584)
US-09-721-480-6

Query Match	100.0%;	Score 5882;	DB 4;	Length 5882;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5882;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCGCGCGTTTCGGTGATGACGGTGAAAACTCTGACACATGACAGCTCCCGGAGACGGTCA	60	
DB	1	TCGCGCGTTTCGGTGATGACGGTGAAAACTCTGACACATGACAGCTCCCGGAGACGGTCA	60	
QY	61	CAGCTTGCTGTGAAGCGGATGCGGGAGCAGACACAGCCGTCAGGCGGCTCAGCGGGTG	120	
DB	61	CAGCTTGCTGTGAAGCGGATGCGGGAGCAGACACAGCCGTCAGGCGGCTCAGCGGGTG	120	
QY	121	TTGGCGGGTGTCGGGGCTGGCTTAACATATCGCGGCATCAGACAGATTGTACTGAGAGTGC	180	
DB	121	TTGGCGGGTGTCGGGGCTGGCTTAACATATCGCGGCATCAGACAGATTGTACTGAGAGTGC	180	
QY	181	ACCATATGAAGCTTTTTCAGAAAGCTAGGCTCCAAAAAGCCTCCTCACTACTTCTGG	240	
DB	181	ACCATATGAAGCTTTTTCAGAAAGCTAGGCTCCAAAAAGCCTCCTCACTACTTCTGG	240	
QY	241	AATAGCTCAGAGCGCGGCGGCTCGGCTCTGCATTAATAAAAAAATTAGTCAGCCA	300	
DB	241	AATAGCTCAGAGCGCGGCGGCTCGGCTCTGCATTAATAAAAAAATTAGTCAGCCA	300	
QY	301	TGGGCGGAGAAATGGGCGGAACTGGGCGGGAGGGAATTATTGGCTATTGGCCATTGCAT	360	

|||||
301 TGGGGCGGAGATGGCGGGAATGGCGGGGAGGGAATATTGGCTATTGGCCATTGCAT 360
QY
|||||
361 AGTTCTATCTATPATCAATAATGTATCATATTTATTTGGCTCATGTCCAATATGACGCCCA 420
Db
|||||
361 AGTTCTATCTATPATCAATAATGTATCATATTTATTTGGCTCATGTCCAATATGACGCCCA 420
QY
|||||
421 TGTGTGACATGATTATGTACTAGTTATTAATAGTAATCAATTACGGGGGTCAATTAGTTTCA 480
Db
|||||
421 TGTGTGACATTTGATTATTTGACTAGTTATTAATAGTAATCAATTACGGGGGTCAATTAGTTTCA 480
QY
|||||
481 AGCCCATATATGGAGTTCGCGCTTACATAAATTTACGGTAAATGGCCCGCTGGCTGACCG 540
Db
|||||
481 AGCCCATATATGGAGTTCGCGCTTACATAAATTTACGGTAAATGGCCCGCTGGCTGACCG 540
QY
|||||
541 CCCAACGACCCCGCCCATTTGACGTCAATAATAGACGTATGTTCCCATAGTAAGCCCAATA 600
Db
|||||
541 CCCAACGACCCCGCCCATTTGACGTCAATAATAGACGTATGTTCCCATAGTAAGCCCAATA 600
QY
|||||
601 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTTGGCAGTA 660
Db
|||||
601 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTTGGCAGTA 660
QY
|||||
661 CATCAAGTGTATCATATGCAATGCGCCCTTATTGACGTCAATGACGCTAAATGGCC 720
Db
|||||
661 CATCAAGTGTATCATATGCAATGCGCCCTTATTGACGTCAATGACGCTAAATGGCC 720
QY
|||||
721 GCGTGCAATATGCCCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC 780
Db
|||||
721 GCGTGCAATATGCCCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC 780
QY
|||||
781 GTATTAGTCAATCGCTATTACCATGTTGATGCGGTTTGGCAGTACACCAATGGGCGTGA 840
Db
|||||
781 GTATTAGTCAATCGCTATTACCATGTTGATGCGGTTTGGCAGTACACCAATGGGCGTGA 840
QY
|||||
841 TAGCGGTTTGGACTCAGCGGATTTCCAAAGTCTCCACCCTTACGCTCAATGGGAGTTTG 900
Db
|||||
841 TAGCGGTTTGGACTCAGCGGATTTCCAAAGTCTCCACCCTTACGCTCAATGGGAGTTTG 900
QY
|||||
901 TTTTGGCAACAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCGCCCGTTGACG 960
Db
|||||
901 TTTTGGCAACAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCGCCCGTTGACG 960
QY
|||||
961 CAAATGGGCGGTAGGCGTCTACGGTGGGAGTCTATATAGCAGAGCTCGTTTAGTGAAC 1020
Db
|||||
961 CAAATGGGCGGTAGGCGTCTACGGTGGGAGTCTATATAGCAGAGCTCGTTTAGTGAAC 1020
QY
|||||
1021 CGTCAGATCGCTTGAGAGCGCCATCCACGCTGTTTTTGACCTCCATAGAAGACACCGGGAC 1080
Db
|||||
1021 CGTCAGATCGCTTGAGAGCGCCATCCACGCTGTTTTTGACCTCCATAGAAGACACCGGGAC 1080
QY
|||||
1081 CGATCCAGCTCCGCGCGCGGAAACGGTGCAATTTGAAACCGGATTCGCCGTGCAAGAGT 1140
Db
|||||
1081 CGATCCAGCTCCGCGCGCGGAAACGGTGCAATTTGAAACCGGATTCGCCGTGCAAGAGT 1140
QY
|||||
1141 GAGGTAAAGTACCGCTATAGACTATAGCCACACCCCTTTGGCTCTTATGATGCTATA 1200
Db
|||||
1141 GAGGTAAAGTACCGCTATAGACTATAGCCACACCCCTTTGGCTCTTATGATGCTATA 1200
QY
|||||
1201 CTGTTTTTGGCTTGGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT 1260
Db
|||||
1201 CTGTTTTTGGCTTGGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT 1260
QY
|||||
1261 TAGCCCTATAGGTGTTGGTTATTGACCAATTTAGCACTCCCTATTTGGTGAAGTACTT 1320
Db
|||||
1261 TAGCCCTATAGGTGTTGGTTATTGACCAATTTAGCACTCCCTATTTGGTGAAGTACTT 1320
QY
|||||
1321 TCCATTACTAATCCATAAATGGCTTTTGGCAACTATCTATTTGGCTATATGCCAA 1380
Db
|||||
1321 TCCATTACTAATCCATAAATGGCTTTTGGCAACTATCTATTTGGCTATATGCCAA 1380
QY
|||||
1381 TACTCTGTCTTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTAT 1440

1381 TACTCTGTCTTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTAT 1440
QY
|||||
1441 TATTTACAAATTTCAATATACAAACGCGCTCCCGTCCCGCGCAGTTTATTATTAACA 1500
Db
|||||
1441 TATTTACAAATTTCAATATACAAACGCGCTCCCGTCCCGCGCAGTTTATTATTAACA 1500
QY
|||||
1501 TAGCGTGGGATCTCCGACATCTCGGTACGTGTTCCGGACATGGGCTCTTCTCCGGTAGC 1560
Db
|||||
1501 TAGCGTGGGATCTCCGACATCTCGGTACGTGTTCCGGACATGGGCTCTTCTCCGGTAGC 1560
QY
|||||
1561 GCGGAGCTTCCACATCTCCGAGCCCTCGTCCCATCCGTCACAGCGGCTCATGGTCCGTCGGC 1620
Db
|||||
1561 GCGGAGCTTCCACATCTCCGAGCCCTCGTCCCATCCGTCACAGCGGCTCATGGTCCGTCGGC 1620
QY
|||||
1621 AGCTCTTCTCTTAAAGTGGAGGCGACATTTAGGCAACAGCAATGCCACCAACACC 1680
Db
|||||
1621 AGCTCTTCTCTTAAAGTGGAGGCGACATTTAGGCAACAGCAATGCCACCAACACC 1680
QY
|||||
1681 AGTGTGCCCAACAGGCCGCTGGGTAGGTATGCTCTGAAAATGAGCTCGGAGATTGG 1740
Db
|||||
1681 AGTGTGCCCAACAGGCCGCTGGGTAGGTATGCTCTGAAAATGAGCTCGGAGATTGG 1740
QY
|||||
1741 GCTCGCACCTGACCGCAGATGGAAGACTTTAAGGCGAGCGGCGAGAAAGATGCAGGCAGCT 1800
Db
|||||
1741 GCTCGCACCTGACCGCAGATGGAAGACTTTAAGGCGAGCGGCGAGAAAGATGCAGGCAGCT 1800
QY
|||||
1801 GAGTTGTTGTAATTCGATAAGAGTACAGGTAATCTCCGTTGGGTGCTGTTAAACGGTGG 1860
Db
|||||
1801 GAGTTGTTGTAATTCGATAAGAGTACAGGTAATCTCCGTTGGGTGCTGTTAAACGGTGG 1860
QY
|||||
1861 AGGGCAGTGTACTGTGAGCAGTACTCGTTGCTGCGCGCGCCGCCACACGACATAATAGCT 1920
Db
|||||
1861 AGGGCAGTGTACTGTGAGCAGTACTCGTTGCTGCGCGCGCCGCCACACGACATAATAGCT 1920
QY
|||||
1921 GACAGACTAACAGACTGTTCCCTTTCCATGGGCTTTTTCTGCAAGTACCCGTCGTCGACGAA 1980
Db
|||||
1921 GACAGACTAACAGACTGTTCCCTTTCCATGGGCTTTTTCTGCAAGTACCCGTCGTCGACGAA 1980
QY
|||||
1981 TTCAGCAATCATGTGATGCAATGAAAGAGGCTCTGCTGTGCTGCTGCTGTGTGGAG 2040
Db
|||||
1981 TTCAGCAATCATGTGATGCAATGAAAGAGGCTCTGCTGTGCTGCTGCTGTGTGGAG 2040
QY
|||||
2041 CAGTCTTCGTTTCGCCGCGCTAGCGAAACCCACGTCACCGGGGGAAGTCCCGGCCACA 2100
Db
|||||
2041 CAGTCTTCGTTTCGCCGCGCTAGCGAAACCCACGTCACCGGGGGAAGTCCCGGCCACA 2100
QY
|||||
2101 CTGTGTCTGGATTGTTAGCTCTCTCGCACAGGCGCCAAAGCAGAACGTCAGCTGATCA 2160
Db
|||||
2101 CTGTGTCTGGATTGTTAGCTCTCTCGCACAGGCGCCAAAGCAGAACGTCAGCTGATCA 2160
QY
|||||
2161 ACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACCTGCATGATAGCTCAACA 2220
Db
|||||
2161 ACACCAACGGCAGTTGGCACCTCAATAGCACGGCCCTGAACCTGCATGATAGCTCAACA 2220
QY
|||||
2221 CCGGCTGGTTGTCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCCTGAGA 2280
Db
|||||
2221 CCGGCTGGTTGTCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCCTGAGA 2280
QY
|||||
2281 GGCTAGCAGCTGCGGACCCCTTACCGATTTTGAACAGGGCTGGGGCCCTATCAGTTATG 2340
Db
|||||
2281 GGCTAGCAGCTGCGGACCCCTTACCGATTTTGAACAGGGCTGGGGCCCTATCAGTTATG 2340
QY
|||||
2341 CCAAAGGAAGCGGCCCGGACGAGCGCCCTACTGTGGCACTACCCCCCAAAACCTTGGC 2400
Db
|||||
2341 CCAAAGGAAGCGGCCCGGACGAGCGCCCTACTGTGGCACTACCCCCCAAAACCTTGGC 2400
QY
|||||
2401 GTATTGTGCCCGGAAGAGTGTGTGTCGGTATATTGCTTCACTCCAGCCCCCGTGG 2460
Db
|||||
2401 GTATTGTGCCCGGAAGAGTGTGTGTCGGTATATTGCTTCACTCCAGCCCCCGTGG 2460
QY
|||||
2461 TGGTGGGAACGACCGCAGGTGCGGCGGCCCACTTACAGCTGGGTGAAATGATACGG 2520
Db
|||||
2461 TGGTGGGAACGACCGCAGGTGCGGCGGCCCACTTACAGCTGGGTGAAATGATACGG 2520

QY 2521 ACCTCTTCGTCCTTAAACAATACAGGCCACCGCTGGGCAATATGGTTCCGGTTGTACCTGGA 2580
DB |||||
QY 2521 ACCTCTTCGTCCTTAAACAATACAGGCCACCGCTGGGCAATATGGTTCCGGTTGTACCTGGA 2580
DB |||||
QY 2581 TGAATCAACTGATGATACCAAAAGTGTGCGGAGCGCTCTCTGTGTCAATCGGAGGGGCGG 2640
DB |||||
QY 2581 TGAATCAACTGATGATACCAAAAGTGTGCGGAGCGCTCTCTGTGTCAATCGGAGGGGCGG 2640
DB |||||
QY 2641 GCAACAACCCCTGCACTGCCCACTGATGCTTCCGCAAGCATCCGGAGCCACATACT 2700
DB |||||
QY 2641 GCAACAACCCCTGCACTGCCCACTGATGCTTCCGCAAGCATCCGGAGCCACATACT 2700
DB |||||
QY 2701 CTCGGTCGGCTCCGGTCCCTGGATCACACCCAGGTGCTTGGTCACTACCCGATAGGC 2760
DB |||||
QY 2701 CTCGGTCGGCTCCGGTCCCTGGATCACACCCAGGTGCTTGGTCACTACCCGATAGGC 2760
DB |||||
QY 2761 TTTGGCATATCTTTGATACATCAACATCAACATATTTAAATCAAGATGTAAGTGGAG 2820
DB |||||
QY 2761 TTTGGCATATCTTTGATACATCAACATCAACATATTTAAATCAAGATGTAAGTGGAG 2820
DB |||||
QY 2821 GGCTCGAACAAGCTGGAAGCTCCCTGCAACTGGACGCGGGGCAACGTTGCGATCTGG 2880
DB |||||
QY 2821 GGCTCGAACAAGCTGGAAGCTCCCTGCAACTGGACGCGGGGCAACGTTGCGATCTGG 2880
DB |||||
QY 2881 AAGATAGGACAGGTCGAGATCGATATGAGAACATCAATCAGGATTCCTTAGGACCCC 2940
DB |||||
QY 2881 AAGATAGGACAGGTCGAGATCGATATGAGAACATCAATCAGGATTCCTTAGGACCCC 2940
DB |||||
QY 2941 TGCTCGTTTACAGCGGGGTTTTCTTTGTGTGAAGAATTCCTCAATAACCCAGAGTC 3000
DB |||||
QY 2941 TGCTCGTTTACAGCGGGGTTTTCTTTGTGTGAAGAATTCCTCAATAACCCAGAGTC 3000
DB |||||
QY 3001 TAGACTCGTGGTGGATCTCTCAATTTCTAGGGGATCTCCGCTGTGTCTTGGCCAA 3060
DB |||||
QY 3001 TAGACTCGTGGTGGATCTCTCAATTTCTAGGGGATCTCCGCTGTGTCTTGGCCAA 3060
DB |||||
QY 3061 ATTTCGAGTCCCAACTCCAACTCACTCAACCACTCTCTCTCTCAATTTGTCTGGTT 3120
DB |||||
QY 3061 ATTTCGAGTCCCAACTCCAACTCACTCAACCACTCTCTCTCTCAATTTGTCTGGTT 3120
DB |||||
QY 3121 ATCGCTGGATGTGTGCGCGGTTTTATCAATATTCCTTCAATCTCTGCTGTATGCTCA 3180
DB |||||
QY 3121 ATCGCTGGATGTGTGCGCGGTTTTATCAATATTCCTTCAATCTCTGCTGTATGCTCA 3180
DB |||||
QY 3181 TCTTCTATTTGGTCTCTCGATATCAAGGTATGTGCGGTTTTGCTCTCAATTTCCAG 3240
DB |||||
QY 3181 TCTTCTATTTGGTCTCTCGATATCAAGGTATGTGCGGTTTTGCTCTCAATTTCCAG 3240
DB |||||
QY 3241 GATCAACAACAACAGTACGGGACCAATGCAAAACCTGCAAGACTCTCTGCTCAAGGCAACT 3300
DB |||||
QY 3241 GATCAACAACAACAGTACGGGACCAATGCAAAACCTGCAAGACTCTCTGCTCAAGGCAACT 3300
DB |||||
QY 3301 CTATGTTTCCCTCATGTGCTGTACAAAACCTACGAGTGAATATGCACCTGTATTCCCA 3360
DB |||||
QY 3301 CTATGTTTCCCTCATGTGCTGTACAAAACCTACGAGTGAATATGCACCTGTATTCCCA 3360
DB |||||
QY 3361 TCCCATCGTCTGGGCTTTCCGAAAATACCTATGAGGATGGGCTCAGTCCGTTCTCTT 3420
DB |||||
QY 3361 TCCCATCGTCTGGGCTTTCCGAAAATACCTATGAGGATGGGCTCAGTCCGTTCTCTT 3420
DB |||||
QY 3421 GGCTCAGTTTACTAGTGCCATTTGTTCAAGTGTTCGTAGGGCTTTCCCCACATGTTTGGC 3480
DB |||||
QY 3421 GGCTCAGTTTACTAGTGCCATTTGTTCAAGTGTTCGTAGGGCTTTCCCCACATGTTTGGC 3480
DB |||||
QY 3481 TTTTACGCTATATGGAATGATGTGATTTGGGGGCAAGTCTGTACAGCATCGTGAAGTCCCT 3540
DB |||||
QY 3481 TTTTACGCTATATGGAATGATGTGATTTGGGGGCAAGTCTGTACAGCATCGTGAAGTCCCT 3540
DB |||||
QY 3541 TTATACGCTGTTACCAATTTTCTTTGCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
DB |||||
QY 3541 TTATACGCTGTTACCAATTTTCTTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
DB |||||

QY 3601 AGCAAGTCTAGAAAGCGCGCCAAAGATATCAAGATCCACTACGCGTTAGAGTCGCTGA 3660
DB |||||
QY 3601 AGCAAGTCTAGAAAGCGCGCCAAAGATATCAAGATCCACTACGCGTTAGAGTCGCTGA 3660
DB |||||
QY 3661 TCAGCCTCGACTGTGTGCTTCTAGTTGCCAGCATCTGTGTTGTCGCCCTCCCGCGTGCCT 3720
DB |||||
QY 3661 TCAGCCTCGACTGTGTGCTTCTAGTTGCCAGCATCTGTGTTGTCGCCCTCCCGCGTGCCT 3720
DB |||||
QY 3721 TCCTTCACCTCGAAGGTGCCACTCCACTGTCTTTCTTAATAAATAGAGAAATGCA 3780
DB |||||
QY 3721 TCCTTCACCTCGAAGGTGCCACTCCACTGTCTTTCTTAATAAATAGAGAAATGCA 3780
DB |||||
QY 3781 TCGCATTGTCTGATAGGTGTCTATTCTATTCTGGGGGTGGGTGGGGCAGGACAGCAAG 3840
DB |||||
QY 3781 TCGCATTGTCTGATAGGTGTCTATTCTATTCTGGGGGTGGGTGGGGCAGGACAGCAAG 3840
DB |||||
QY 3841 GGGGAGGATTTGGGAAGACATAGCAGGCATGCTGGGAGGCTCTTCGCTTCCTCGCTCAC 3900
DB |||||
QY 3841 GGGGAGGATTTGGGAAGACATAGCAGGCATGCTGGGAGGCTCTTCGCTTCCTCGCTCAC 3900
DB |||||
QY 3901 TGAATCGCTCGCTCGCTCGTTTCGGGTGCGGCGAGCGGTATCAGCTCAGCTCAAAAGCGGT 3960
DB |||||
QY 3901 TGAATCGCTCGCTCGCTCGTTTCGGGTGCGGCGAGCGGTATCAGCTCAGCTCAAAAGCGGT 3960
DB |||||
QY 3961 AATACGCTTTATCCACAGAATCAGGGGATAAACGAGGAAAGAACATGTAGCAAAAGSCCA 4020
DB |||||
QY 3961 AATACGCTTTATCCACAGAATCAGGGGATAAACGAGGAAAGAACATGTAGCAAAAGSCCA 4020
DB |||||
QY 4021 GCMAAGGCGAGAACCGTAAAGGCGCGGTGCTGGCGTTTTTTCATAGGCTCCGCC 4080
DB |||||
QY 4021 GCMAAGGCGAGAACCGTAAAGGCGCGGTGCTGGCGTTTTTTCATAGGCTCCGCC 4080
DB |||||
QY 4081 CCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACT 4140
DB |||||
QY 4081 CCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACT 4140
DB |||||
QY 4141 ATAAAGATACAGGCGTTTTCCCTCGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCCT 4200
DB |||||
QY 4141 ATAAAGATACAGGCGTTTTCCCTCGGAAGCTCCCTCGTGCCTCTCTGTTCGACCCCT 4200
DB |||||
QY 4201 GCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATG 4260
DB |||||
QY 4201 GCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCGTGGCGCTTCTCATG 4260
DB |||||
QY 4261 CTCACGCTGTAGTATCTCAGTTTCGCTGTAGTGTGCTCCAGTGGGCTGTGTGCA 4320
DB |||||
QY 4261 CTCACGCTGTAGTATCTCAGTTTCGCTGTAGTGTGCTCCAGTGGGCTGTGTGCA 4320
DB |||||
QY 4321 CGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACCTATCGTCTGAGTCCAA 4380
DB |||||
QY 4321 CGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACCTATCGTCTGAGTCCAA 4380
DB |||||
QY 4381 CCGGTAAAGACAGCTTATCGGCACTGGCAGCAGCAGCTGTGTAAAGATAGCAGAGC 4440
DB |||||
QY 4381 CCGGTAAAGACAGCTTATCGGCACTGGCAGCAGCAGCTGTGTAAAGATAGCAGAGC 4440
DB |||||
QY 4441 GAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGCGCTTAACTACGCTACACTAG 4500
DB |||||
QY 4441 GAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGCGCTTAACTACGCTACACTAG 4500
DB |||||
QY 4501 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGG 4560
DB |||||
QY 4501 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGG 4560
DB |||||
QY 4561 TAGCTCTTGTATCCGGCAACAAACCCGCTGTAGCGGTGGTTTTTGTGTTGCAAGCA 4620
DB |||||
QY 4561 TAGCTCTTGTATCCGGCAACAAACCCGCTGTAGCGGTGGTTTTTGTGTTGCAAGCA 4620
DB |||||
QY 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTCTTTCTACGGGTC 4680
DB |||||
QY 4621 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGTCTTTCTACGGGTC 4680
DB |||||
QY 4681 TGACGCTCAGTGGAAAGAACTCAGCTTAAAGGATTTTGGTCTAGATTTATCAAAAG 4740

|||||
4681 TGACGCTAGTGGACGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTAACAAG 4740
|||||
4741 GATCTTCACCTAGATCCTTTTAAATTAATAAGTATTAATCAATCTAAAGTATATA 4800
|||||
4741 GATCTTCACCTAGATCCTTTTAAATTAATAAGTATTAATCAATCTAAAGTATATA 4800
|||||
4801 TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAGGCAACCTATCTCAGCGAT 4860
|||||
4801 TGAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAGGCAACCTATCTCAGCGAT 4860
|||||
4861 CTGCTCTATTTCTGTTATCATCATAGTTGCTCGTAGCTCCCGTCTGTGATGATTAACATACG 4920
|||||
4861 CTGCTCTATTTCTGTTATCATCATAGTTGCTCGTAGCTCCCGTCTGTGATGATTAACATACG 4920
|||||
4921 GGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCAGC 4980
|||||
4921 GGAGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGGAGACCCACGCTCACCAGC 4980
|||||
4981 TCAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGGAGCGGAGAGTGGTCTCTGC 5040
|||||
4981 TCAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGGAGCGGAGAGTGGTCTCTGC 5040
|||||
5041 AACTTTATCGCTCCATCCAGTCTATTAATTTGTTGCCGGAAGCTAGAGTAAGTAGTTC 5100
|||||
5041 AACTTTATCGCTCCATCCAGTCTATTAATTTGTTGCCGGAAGCTAGAGTAAGTAGTTC 5100
|||||
5101 GCCAGTTAATAGTTTGGCAAGCTGTTGCCATGCTACAGGCATCGTGGTGTACGCTC 5160
|||||
5101 GCCAGTTAATAGTTTGGCAAGCTGTTGCCATGCTACAGGCATCGTGGTGTACGCTC 5160
|||||
5161 GTGCTTTGATAGGCTTCAATCAGCTCCGGTTCCTCCAAAGATCAAGCGAGTTACATGATC 5220
|||||
5161 GTGCTTTGATAGGCTTCAATCAGCTCCGGTTCCTCCAAAGATCAAGCGAGTTACATGATC 5220
|||||
5221 CCCATGTTGTGCAAAAGCGTTAGCTCTTCCGTCTCCGATCGTTGTGAGAGTAA 5280
|||||
5221 CCCATGTTGTGCAAAAGCGTTAGCTCTTCCGTCTCCGATCGTTGTGAGAGTAA 5280
|||||
5281 GTTGGCCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATTCCTTACTGTCTAT 5340
|||||
5281 GTTGGCCGAGTGTATCACTCATGTTATGCGAGCACTGCATAATTCCTTACTGTCTAT 5340
|||||
5341 GCCATCCGTAAGTATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTGAGAATA 5400
|||||
5341 GCCATCCGTAAGTATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTGAGAATA 5400
|||||
5401 GTGTATGCGGACCGAGTTGCTTCTGCGCGGTCAATACGGGATAATACCGCGCCACA 5460
|||||
5401 GTGTATGCGGACCGAGTTGCTTCTGCGCGGTCAATACGGGATAATACCGCGCCACA 5460
|||||
5461 TAGCAGAACTTTAAAGTCTCATCATTTGGAACCGTTCTTGGGGCGAAAACTCTCAAG 5520
|||||
5461 TAGCAGAACTTTAAAGTCTCATCATTTGGAACCGTTCTTGGGGCGAAAACTCTCAAG 5520
|||||
5521 GATCTTACCGCTGTTGAGATCCAGTTGATGTAACCACTCTGTGACCCCACTGATCTTC 5580
|||||
5521 GATCTTACCGCTGTTGAGATCCAGTTGATGTAACCACTCTGTGACCCCACTGATCTTC 5580
|||||
5581 AGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGC 5640
|||||
5581 AGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGC 5640
|||||
5641 AAAAAAGGAATAAGGGCGACACGGAATGTTGAATACATCTCTCTCTTTTCAATA 5700
|||||
5641 AAAAAAGGAATAAGGGCGACACGGAATGTTGAATACATCTCTCTCTTTTCAATA 5700
|||||
5701 TTATTGAACATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGATTTTA 5760
|||||
5701 TTATTGAACATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGATTTTA 5760
|||||
5761 GAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGCTTA 5820
|||||

Db 5761 GAAAAATAAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTCTA 5820
QY 5821 AGAAACCAATTATTATCATGACATTAACCTATAAAATAGGCGTATACACAGGCCCTTTCCG 5880
|||||
Db 5821 AGAAACCAATTATTATCATGACATTAACCTATAAAATAGGCGTATACACAGGCCCTTTCCG 5880
QY 5881 TC 5882
|||||
Db 5881 TC 5882
|||||
RESULT 2
US-09-721-480-4
; Sequence 4, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PPO1635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCMVII opti
; OTHER INFORMATION: 330 E1/SAG
; NAME/KEY: CDS
; LOCATION: (1992).. (3161)
US-09-721-480-4
Query Match 78.5%; Score 4617.2; DB 4; Length 5459;
Best Local Similarity 89.1%; Pred. No. 0; Mismatches 218; Indels 423; Gaps 7;
Matches 5241; Conservative 0;
QY 1 TCGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
Db 1 TCGCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAACGGATGCCGGGAGCAGACAAGCCGTCAGGCGCGTCAGCGGGTG 120
Db 61 CAGCTTGTCTGTAACGGATGCCGGGAGCAGACAAGCCGTCAGGCGCGTCAGCGGGTG 120
QY 121 TTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
Db 121 TTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
QY 181 ACCATATGAAGCTTTTTCGCAAAAGCTAGGCTCCAAAAAGCCCTCTCCTACTTCTCGG 240
Db 181 ACCATATGAAGCTTTTTCGCAAAAGCTAGGCTCCAAAAAGCCCTCTCCTACTTCTCGG 240
QY 241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAATAATAAAAAAATTAAGTCAGCCA 300
Db 241 AATAGCTCAGAGCCGAGCGGCTCGGCTCTGCATAATAATAAAAAAATTAAGTCAGCCA 300
QY 301 TGGGCGGAGAAATGGCGGAACCTGGCGGGGAGGAAATTTATGGCTATTGGCCATTGCAT 360
Db 301 TGGGCGGAGAAATGGCGGAACCTGGCGGGGAGGAAATTTATGGCTATTGGCCATTGCAT 360
QY 361 AGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACGCCA 420
Db 361 AGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGACGCCA 420
QY 421 TGTGACATTTGATTTAGCTAGTTTATTAATAGTAAATCAATTACGGGGTCATTAGTTTAT 480
Db 421 TGTGACATTTGATTTAGCTAGTTTATTAATAGTAAATCAATTACGGGGTCATTAGTTTAT 480
QY 481 AGCCCATATATGAGGTTTCGCGCTTACATAAATTAACGTTAAATGGCCCGCTGGCTGACCG 540

Db 481 AGCCCATATATGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCG 540
Qy 541 CCCAAGACCCCGCCCATAGCTCAATAATGAAGTATGTTCCCATAGTACGCCAATA 600
Db 541 CCCAAGACCCCGCCCATAGCTCAATAATGAAGTATGTTCCCATAGTACGCCAATA 600
Qy 601 GGGACTTTTCATGAGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA 660
Db 601 GGGACTTTTCATGAGCTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA 660
Qy 661 CATCAAGTGTATCATATGCAAGTCCGCCCCCTATTGAGCTCAATGACGGTAAATGGCCC 720
Db 661 CATCAAGTGTATCATATGCAAGTCCGCCCCCTATTGAGCTCAATGACGGTAAATGGCCC 720
Qy 721 GCTGGCATATGCCCCAGTACATGACTTACGGGACTTTTCCAGGACTTTTCCACTTTGGCAGTACATCTAC 780
Db 721 GCTGGCATATGCCCCAGTACATGACTTACGGGACTTTTCCAGGACTTTTCCACTTTGGCAGTACATCTAC 780
Qy 781 GTATTAGTCAATGCTATTACCATGATGCGGTTTTTGGCAGTACACCAATGGGCGTGA 840
Db 781 GTATTAGTCAATGCTATTACCATGATGCGGTTTTTGGCAGTACACCAATGGGCGTGA 840
Qy 841 TAGCGTTTGAATCAGCGGATTTCCAAAGTCTCCACCCCAATTGACGTCAATGGGAGTTTG 900
Db 841 TAGCGTTTGAATCAGCGGATTTCCAAAGTCTCCACCCCAATTGACGTCAATGGGAGTTTG 900
Qy 901 TTTTGGCACCANAATCAACGGGACTTTCCANAATGTCGTAATAACCCCGCCCGTTGACG 960
Db 901 TTTTGGCACCANAATCAACGGGACTTTCCANAATGTCGTAATAACCCCGCCCGTTGACG 960
Qy 961 CAAATGGGCGTAGGGCTGACGGTGGAGTCTATATAGCAGAGCTCGTTTGTAGTGAAC 1020
Db 961 CAAATGGGCGTAGGGCTGACGGTGGAGTCTATATAGCAGAGCTCGTTTGTAGTGAAC 1020
Qy 1021 CGTCAGATCGCTTGAGAGCCCATCAACGCTGTTTGAACCTCCATAGAAGACACCGGGAC 1080
Db 1021 CGTCAGATCGCTTGAGAGCCCATCAACGCTGTTTGAACCTCCATAGAAGACACCGGGAC 1080
Qy 1081 CGATCAGCTCGCGCGGGGACCGGTGATGGAACGGGATTTCCCGTGGCAGAGT 1140
Db 1081 CGATCAGCTCGCGCGGGGACCGGTGATGGAACGGGATTTCCCGTGGCAGAGT 1140
Qy 1141 GAGTAAAGTACCGCTATAGACTATAGGCACACCCCTTTGGCTCTTATGATGCTATA 1200
Db 1141 GAGTAAAGTACCGCTATAGACTATAGGCACACCCCTTTGGCTCTTATGATGCTATA 1200
Qy 1201 CTGTTTTTGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT 1260
Db 1201 CTGTTTTTGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGATAGCT 1260
Qy 1261 TAGCCTATAGGTGGGTTATGACCATTTATGACCACTCCCTATTTGGTACGATATT 1320
Db 1261 TAGCCTATAGGTGGGTTATGACCATTTATGACCACTCCCTATTTGGTACGATATT 1320
Qy 1321 TCCATTACTAAATCCATAAATGCTCTTTGCCAACAATCTCTATTTGGTATATGCCAA 1380
Db 1321 TCCATTACTAAATCCATAAATGCTCTTTGCCAACAATCTCTATTTGGTATATGCCAA 1380
Qy 1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGCTCAATTTAT 1440
Db 1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGCTCAATTTAT 1440
Qy 1441 TATTTACAAATTCATATACAAACAGCGCTCCCGGTGCGGAGTTTATTTAAACA 1500
Db 1441 TATTTACAAATTCATATACAAACAGCGCTCCCGGTGCGGAGTTTATTTAAACA 1500
Qy 1501 TAGCGTGGGATCTCCGACATCTCGGTTAGGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
Db 1501 TAGCGTGGGATCTCCGACATCTCGGTTAGGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
Qy 1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCCAGCGGCTCATGGTGGCTCGGC 1620

Db 1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCCAGCGGCTCATGGTGGCTCGGC 1620
Qy 1621 AGCTCCTTCTCTCTTAACAGTGGAGCCAGACTTAGGCACAGCAATGCCACACACACC 1680
Db 1621 AGCTCCTTCTCTCTTAACAGTGGAGCCAGACTTAGGCACAGCAATGCCACACACACC 1680
Qy 1681 AGTGTGCCGCAACAAGCCCGTGGCGGTAGGTATGTCTGAAATGAGCTCGAGATTGG 1740
Db 1681 AGTGTGCCGCAACAAGCCCGTGGCGGTAGGTATGTCTGAAATGAGCTCGAGATTGG 1740
Qy 1741 GCTCGCACCTGACGCGCAGATGAAGACTTAAGCAGCGGCGAGAGATGCGGCGAGCT 1800
Db 1741 GCTCGCACCTGACGCGCAGATGAAGACTTAAGCAGCGGCGAGAGATGCGGCGAGCT 1800
Qy 1801 GAGTTGTCTGATTTCTGATAAGAGTACAGAGTAACTCCCGTTGGCGTCTGTAAACGGTGG 1860
Db 1801 GAGTTGTCTGATTTCTGATAAGAGTACAGAGTAACTCCCGTTGGCGTCTGTAAACGGTGG 1860
Qy 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCCGCGCCACACAGACATAATAGCT 1920
Db 1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCCGCGCCACACAGACATAATAGCT 1920
Qy 1921 GACAGACTAACAGACTGTTCCCTTTCCATGGGCTCTTTCTGCAGTCAACCGTCTCGACGAA 1980
Db 1921 GACAGACTAACAGACTGTTCCCTTTCCATGGGCTCTTTCTGCAGTCAACCGTCTCGACGAA 1980
Qy 1981 TTTCAAGCAATCATGATGCAATGAAGAGAGGCTCTGCTGTGTGCTGCTGTGTGGAG 2040
Db 1981 TTTCAAGCAATCATGATGCAATGAAGAGAGGCTCTGCTGTGTGCTGCTGTGTGGAG 2040
Qy 2041 CAGTCTTCTGTTTGGCCAGCGCTAGCGAAACCCACGCTCACCGGGGGAAGTGCCTGGCCACA 2100
Db 2041 CAGTCTTCTGTTTGGCCAGCGCTAGCTA----- 2068
Qy 2101 CTGTGTCTGGATTTGTAGCTCTCTCGACACAGCGCCGCAAGAGACGTCAGCTGATCA 2160
Db 2069 ----- 2079
Qy 2161 ACACCAACGCGAGTTGGCACCTCAATAGCACGCGCCCTGAACATGATAGCTCAACA 2220
Db 2080 ACAGCACCGGCTGTACACAGTGA-----CCAAGACTGCCCCNACA 2121
Qy 2221 CCGGCTGGTGGCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGGTGTCTTGAGA 2280
Db 2122 GCAGCATCTGTGTACGAGGCGCGCGCATCTGCACACCCCGCTGCGTGC----- 2176
Qy 2281 GGTAGCAGCTGCGGACCCCTTACCGATTTTGAACAGGGCTGGGGCTTATCAGTTATG 2340
Db 2177 -----CTGCGTGGCGAGGGCAACGCGCGCTGCTGGGTGGCCATGACCCCA 2226
Qy 2341 CCAACGGAAGCGGCGCCGACACGCGCCCTACTGTGGCACTACCCCCCAAAACCTTGGC 2400
Db 2227 CCGTGCCACCCCGGACGCAAGCTGCCCGCCACCCAGCTGCGGCC----- 2273
Qy 2401 GTATTGTGCCCGAAGAGTGTGTGTGTCGGTATATTGCTTCACTCCAGCGCCGTGG 2460
Db 2274 -----CACATCGACCTGC 2286
Qy 2461 TGGTGGGAACGACAGAGTGGGCGCGCCACCTACAGCTGGGTGAAATGATACGG 2520
Db 2287 TGGTGGGACGCGCACCTGTGTGAGCGCC-----TGT 2319
Qy 2521 AGCTCTTCTCTCTTAACATACAGCGCACCGCTGGGCAATTTGGTTGTTGTTACCTGGA 2580
Db 2320 AGTGGGCGACCTGTGGCGAGCGTTCCTGTGGGCGCAGCTGTTTC----- 2366
Qy 2581 TGAACCTCAACTGGATTTCAACAAAGTGTGGGAGCGCTCTCTGTGTATCATCGAGGGGCGG 2640
Db 2367 -----ACCT 2370
Qy 2641 GAAACAAACCTGTGACCTGCCACATGATTTGTTTCCGGAAGCATCCGAGCCACATACT 2700
Db 2371 TCAGCCCCCGCGCCACTGGAACCAACCGGGCTGCAACTGCAGCATCTACCCCGGCACA 2430

QY	2701	CTCGGTGGGCTCCGGTCCCTGATCACAACCCAGGTGCCTGGTCGACTACCCGTATAGGC	2760
DB	2431	TCACCGGCCACCGCATGGCTGGGACATGATGAT-----	2464
QY	2761	TTTGGCATATCCTTGTFACCATCAACTACACCATATTTAAATCAGGATGTAGTGGGAG	2820
DB	2465	-----	2464
QY	2821	GGGTGGAACACAGGCTGGAAGTGCCTGCAACTGACGGGGCGAACTGTGCGATCTGG	2880
DB	2465	-----	2464
QY	2881	AGATAGGACAGGTCCGAGATCGATATGGAGAACATCATCATCAGGATTCCTAGGACCCC	2940
DB	2465	-----GAACTGGAGCCCCACCAACCATGGAGAACATCATCAGGATTCCTAGGACCCC	2517
QY	2941	TGCTCGTGTACAGGGGGGTTTTCTGTGTGACAGAAATCCTCACAATACCGCAGAGTC	3000
DB	2518	TGCTCGTGTACAGGGGGGTTTTCTGTGTGACAGAAATCCTCACAATACCGCAGAGTC	2577
QY	3001	TAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGATCTCCCGTGTGTCTTTGGCCAAA	3060
DB	2578	TAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGATCTCCCGTGTGTCTTTGGCCAAA	2637
QY	3061	ATTTCGAGTCCCAACCTCCAAATCATCAACCACTCCTGTCTCCAAATTTGTCTGGTTT	3120
DB	2638	ATTTCGAGTCCCAACCTCCAAATCACTCAACCACTCCTGTCTCCAAATTTGTCTGGTTT	2697
QY	3121	ATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTCTCATCTCCTGCTGCTATGCCTCA	3180
DB	2698	ATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTCTCATCTCCTGCTGCTATGCCTCA	2757
QY	3181	TCCTCTTATTGGTTCTCTGGATTATCAAGTATGTTGCCGTTTGTCTCTAAATTCAG	3240
DB	2758	TCCTCTTATTGGTTCTCTGGATTATCAAGTATGTTGCCGTTTGTCTCTAAATTCAG	2817
QY	3241	GATCAACAACAACAGTAGCGGACCAATGCAAAACCTTGACAGACTCTGCTCAAGGCAACT	3300
DB	2818	GATCAACAACAACAGTAGCGGACCAATGCAAAACCTTGACAGACTCTGCTCAAGGCAACT	2877
QY	3301	CTATGTTTCCCTCATGTGCTGTACAAAACCTACGGATGGAATTCACCTGTATTCCCA	3360
DB	2878	CTATGTTTCCCTCATGTGCTGTACAAAACCTACGGATGGAATTCACCTGTATTCCCA	2937
QY	3361	TCCCATCGTCTGGGTTTTGCAAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT	3420
DB	2938	TCCCATCGTCTGGGTTTTGCAAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT	2997
QY	3421	GGCTCAGTTTACTAGTGCCATTTGTTGCTAGTGGTTGTTAGGGCTTTCCCGCATGTTTGGC	3480
DB	2998	GGCTCAGTTTACTAGTGCCATTTGTTGCTAGTGGTTGTTAGGGCTTTCCCGCATGTTTGGC	3057
QY	3481	TTTTCAGCTATGAGTGTGTTGTTGGGGCCAACTGTGTACAGCATCGTGAGTCCCT	3540
DB	3058	TTTTCAGCTATGAGTGTGTTGTTGGGGCCAACTGTGTACAGCATCGTGAGTCCCT	3117
QY	3541	TTATACCGCTGTACCAATTTTCTTTGCTCTGGGTATACATTTAAGAAATCAGACTCG	3600
DB	3118	TTATACCGCTGTACCAATTTTCTTTGCTCTGGGTATACATTTAAGAAATCAGACTCG	3177
QY	3601	AGCAAGTCTAGAAAGGGGCCCAAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA	3660
DB	3178	AGCAAGTCTAGAAAGGGGCCCAAGATATCAAGGATCCACTACGCGTTAGAGCTCGCTGA	3237
QY	3661	TCAGCTCGACTGCTGCTTCTAGTTCGACCACTGTTGTTTGGCCCTCCCGCGGCT	3720
DB	3238	TCAGCTCGACTGCTGCTTCTAGTTCGACCACTGTTGTTTGGCCCTCCCGCGGCT	3297
QY	3721	TCCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTCTTAATAAATGAGGAAATTGCA	3780
DB	3298	TCCTTGACCTGGAAGGTGCCACTCCCACTGTCTTTCTTAATAAATGAGGAAATTGCA	3357

QY	3781	TCGCATTGCTCAGTAGGTGTCAATTCTATTCTGGGGGTGGGGTGGGCGAGCACAGCAAG	3840
DB	3358	TCGCATTGCTCAGTAGGTGTCAATTCTATTCTGGGGGTGGGGTGGGCGAGCACAGCAAG	3417
QY	3841	GGGGAGGATTGGGAAGACAATAGCAGGCACTCTTGGGAGCTCTTCGCTTCTCGCTCAC	3900
DB	3418	GGGGAGGATTGGGAAGACAATAGCAGGCACTCTTGGGAGCTCTTCGCTTCTCGCTCAC	3477
QY	3901	TCACTCGCTCGCTCGGTTCGGTCTGGCGAGCGGTATCAGCTCATCAAAAGCGGT	3960
DB	3478	TCACTCGCTCGCTCGGTTCGGTCTGGCGAGCGGTATCAGCTCATCAAAAGCGGT	3537
QY	3961	ANFACGGTTATCCACAGATCAGGGGATAACCCAGGAAGACATGTGACAAAAGGCCA	4020
DB	3538	ANFACGGTTATCCACAGATCAGGGGATAACCCAGGAAGACATGTGACAAAAGGCCA	3597
QY	4021	GCAAAAGGCCAGAACCGTAAAAAGCCGCTTGTCTGGCGTTTTTTCATAGGCTCCGCC	4080
DB	3598	GCAAAAGGCCAGAACCGTAAAAAGCCGCTTGTCTGGCGTTTTTTCATAGGCTCCGCC	3657
QY	4081	CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAGACT	4140
DB	3658	CCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAAACCCGACAGACT	3717
QY	4141	ATAAAGATACACAGGCGTTTTCCCTCGGAGCTCCCTCGTCCGCTCTCTCTTCCGACCT	4200
DB	3718	ATAAAGATACACAGGCGTTTTCCCTCGGAGCTCCCTCGTCCGCTCTCTCTTCCGACCT	3777
QY	4201	GCGCTTACCGGATACCTGTCCGCTTTCTCCTTCGGGAAGCGTGGCGCTTCTCAATG	4260
DB	3778	GCGCTTACCGGATACCTGTCCGCTTTCTCCTTCGGGAAGCGTGGCGCTTCTCAATG	3837
QY	4261	CTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCCTCAAGCTGGGCTGTGTGA	4320
DB	3838	CTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCCTCAAGCTGGGCTGTGTGA	3897
QY	4321	CGAAACCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTACGTTCTGAGTCCAA	4380
DB	3898	CGAAACCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTACGTTCTGAGTCCAA	3957
QY	4381	CCCGGTAAACACGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGC	4440
DB	3958	CCCGGTAAACACGACTTATCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGC	4017
QY	4441	GAGGTATGAGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTTAACCTACGCTACCTAG	4500
DB	4018	GAGGTATGAGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTTAACCTACGCTACCTAG	4077
QY	4501	RAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGGTGG	4560
DB	4078	RAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAGGTGG	4137
QY	4561	TAGCTCTTCTGATCCGGCAAAACCAACCCCTGGTAGCGGTGGTTTTTTTGTTCGAAGCA	4620
DB	4138	TAGCTCTTCTGATCCGGCAAAACCAACCCCTGGTAGCGGTGGTTTTTTTGTTCGAAGCA	4197
QY	4621	GCAGATTACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTTTTCTACGGGGTC	4680
DB	4198	GCAGATTACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTTTTCTACGGGGTC	4257
QY	4681	TGACGCTCAGTGGAACGAAACTCAGTTAAGGGAATTTTGGTCAAGAGATTAACAAGAG	4740
DB	4258	TGACGCTCAGTGGAACGAAACTCAGTTAAGGGAATTTTGGTCAAGAGATTAACAAGAG	4317
QY	4741	GATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCTAAGATATATA	4800
DB	4318	GATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCTAAGATATATA	4377
QY	4801	TGAGTAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGGAT	4860
DB	4378	TGAGTAACTTGGTGTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGGAT	4437
QY	4861	CTGTCTATTCTTCATCCTAGTTGCTCCCGTGGTGTAGATAACTACGATACG	4920

```
Db 4438 CTGCTATTATTCGTTTCCATAGTTCCTGACTCCCGCTGCTAGATAACTACGATACG 4497
Qy 4921 GGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGGAGACCCAGCTCACCGGC 4980
Db 4498 GGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGGAGACCCAGCTCACCGGC 4557
Qy 4981 TCAGATTTATCAGCAATAAACAGCAGCAGCGGAAGGCGGAGAGAGTGTCTCTGC 5040
Db 4558 TCAGATTTATCAGCAATAAACAGCAGCAGCGGAAGGCGGAGAGTGTCTCTGC 4617
Qy 5041 AACTTTATCGGCTTCATCAGTCTATTAATTTGTCGCGGAAGCTAGAGTAAGTAGTTC 5100
Db 4618 AACTTTATCGGCTTCATCAGTCTATTAATTTGTCGCGGAAGCTAGAGTAAGTAGTTC 4677
Qy 5101 GCCAGTTAATAGTTTCGGAACGTTGTCGTCATTTGCTACAGGCATCGTGTGTACGCTC 5160
Db 4678 GCCAGTTAATAGTTTCGGAACGTTGTCGTCATTTGCTACAGGCATCGTGTGTACGCTC 4737
Qy 5161 GTCGTTTGGTATGGCTTCATTCAGTCTCCGTTCCCAACGATCAAGGCGAGTTACATGATC 5220
Db 4738 GTCGTTTGGTATGGCTTCATTCAGTCTCCGTTCCCAACGATCAAGGCGAGTTACATGATC 4797
Qy 5221 CCCATGTTGTGCAAAAAGCGTTAGTCTCTTCGCTCTCCGATCGTTGTTCAGAAAGTAA 5280
Db 4798 CCCATGTTGTGCAAAAAGCGTTAGTCTCTTCGCTCTCCGATCGTTGTTCAGAAAGTAA 4857
Qy 5281 GTTGGCCGAGTGTATCATCTCATGTTATGCGCAGCACTGCATAATTTCTTACTGTCTAT 5340
Db 4858 GTTGGCCGAGTGTATCATCTCATGTTATGCGCAGCACTGCATAATTTCTTACTGTCTAT 4917
Qy 5341 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATTTCTGAGAATA 5400
Db 4918 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATTTCTGAGAATA 4977
Qy 5401 GTGATATGCGGCGACCGAGTGTCTTTGTCGCGGCGTCAATACGGGATAATACCGCGCCACA 5460
Db 4978 GTGATATGCGGCGACCGAGTGTCTTTGTCGCGGCGTCAATACGGGATAATACCGCGCCACA 5037
Qy 5461 TAGCAAACTTTTAAAGTCTCATCTATGGAACGTTCTTGGGCGGAAACTCTCAAG 5520
Db 5038 TAGCAAACTTTTAAAGTCTCATCTATGGAACGTTCTTGGGCGGAAACTCTCAAG 5097
Qy 5521 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCAACTGATCTTC 5580
Db 5098 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACCAACTGATCTTC 5157
Qy 5581 AGCATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGC 5640
Db 5158 AGCATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGC 5217
Qy 5641 AAAAAAGGGAATAAGGCGACCGGAATGTTGAATACTCATCTCTTCTTTTCAATA 5700
Db 5218 AAAAAAGGGAATAAGGCGACCGGAATGTTGAATACTCATCTCTTCTTTTCAATA 5277
Qy 5701 TTATTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTA 5760
Db 5278 TTATTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTA 5337
Qy 5761 GAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGCTTA 5820
Db 5338 GAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGGAAAAAGTGCCACCTGACGCTTA 5397
Qy 5821 AGNAACCATTTATCATGACATTAACCTATAAATAAGCGTATCAAGGCGGCTTTCCG 5880
Db 5398 AGNAACCATTTATCATGACATTAACCTATAAATAAGCGTATCAAGGCGGCTTTCCG 5457
Qy 5881 TC 5882
Db 5458 TC 5459
```

RESULT 3

```
US-09-721-480-2
; Sequence 2, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pCMVII-ps2-Sag
; NAME/KEY: CDS
; LOCATION: (1988)..(2830)
;
US-09-721-480-2
```

```
Query Match 71.2%; Score 4190; DB 4; Length 5128;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 5038; Conservative 0; Mismatches 90; Indels 754; Gaps 4;
```

```
Qy 1 TCGCGGTTTCGGTGTAGCGGTGAAAACTCTGACACATGACGCTCCCGGAGACGGTCA 60
Db 1 TCGCGGTTTCGGTGTAGCGGTGAAAACTCTGACACATGACGCTCCCGGAGACGGTCA 60
Qy 61 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCGTCAGGCGGCTCAGCGGGTG 120
Db 61 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCGTCAGGCGGCTCAGCGGGTG 120
Qy 121 TTGGCGGGTGTCTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGATGAGAGTGC 180
Db 121 TTGGCGGGTGTCTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGATGAGAGTGC 180
Qy 181 ACCATATGAGCTTTTGCAAAAGCCTAGGCTCCAAAAAGCCTCTCACTACTTCTGG 240
Db 181 ACCATATGAGCTTTTGCAAAAGCCTAGGCTCCAAAAAGCCTCTCACTACTTCTGG 240
Qy 241 AATAGCTCAGAGCGCGAGCGGCTCGGCTCTGCATAAATAAATAAATAAATAAATAA 300
Db 241 AATAGCTCAGAGCGCGAGCGGCTCGGCTCTGCATAAATAAATAAATAAATAAATAA 300
Qy 301 TGGGCGGAGAAATGGCGGAACCTGGGCGGGAGGGAATTTATGGCTATTGGCCATTGCAT 360
Db 301 TGGGCGGAGAAATGGCGGAACCTGGGCGGGAGGGAATTTATGGCTATTGGCCATTGCAT 360
Qy 361 AGTTGTATCTATATCATATAATATGATATTTATTTGGCTCATGTCTCAATATGACGCCA 420
Db 361 AGTTGTATCTATATCATATAATATGATATTTATTTGGCTCATGTCTCAATATGACGCCA 420
Qy 421 TGTTCAGATTGATTATTGACTAGTTATTAAATAGTAATCAATTACGGGCTCATTTAGTTTCAT 480
Db 421 TGTTCAGATTGATTATTGACTAGTTATTAAATAGTAATCAATTACGGGCTCATTTAGTTTCAT 480
Qy 481 AGCCCATATATGAGTTCGCGCTTACATAAATTAACGGTAAATGGCCCGCTGGCTGACCG 540
Db 481 AGCCCATATATGAGTTCGCGCTTACATAAATTAACGGTAAATGGCCCGCTGGCTGACCG 540
Qy 541 CCCAAGCAGCCCGCCCATTTGACGTCATAATGACGATATTTGCCATAGTAAGCCCAATA 600
Db 541 CCCAAGCAGCCCGCCCATTTGACGTCATAATGACGATATTTGCCATAGTAAGCCCAATA 600
Qy 601 GGGACTTTTCCATTGACGTCATCAATGGGTGGAGTATTACCGTAAACTCCCACTTTGGCAGTA 660
Db 601 GGGACTTTTCCATTGACGTCATCAATGGGTGGAGTATTACCGTAAACTCCCACTTTGGCAGTA 660
Qy 661 CATCAAGTGTATCATATGCAAGTCGCGCCCTCTATTGACGTCATTAACGCGTAAATGGCCC 720
```

|||||
661 CATCAAGTGTATCATATGCCAAGTCGCGCCCTTATTGACGTCAATGACGGTAATGGCCC 720
QY
721 GCCTGGCATTTATCCCAAGTACATGACTTTACGGGACTTTCTACTTTGGCAGTACATCTAC 780
Db
721 GCCTGGCATTTATGCCAAGTACATGACTTTACGGGACTTTCTACTTTGGCAGTACATCTAC 780
QY
781 GTATTAGTCATCGCTATTACATGGTGATGCGGTTTGGGAGTACACCAATGGGCGTGA 840
Db
781 GTATTAGTCATCGCTATTACATGGTGATGCGGTTTGGGAGTACACCAATGGGCGTGA 840
QY
841 TAGCGGTTTGACTCAGCGGGATTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900
Db
841 TAGCGGTTTGACTCAGCGGGATTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTG 900
QY
901 TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 960
Db
901 TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 960
QY
961 CAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020
Db
961 CAAATGGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020
QY
1021 CGTCAGATCGCCTGGAGAGCCCATCCACGCTGTTTTGACCTCCATAGAAGACACGGGAC 1080
Db
1021 CGTCAGATCGCCTGGAGAGCCCATCCACGCTGTTTTGACCTCCATAGAAGACACGGGAC 1080
QY
1081 CGATCAGGCTCCCGCGCGGGAAACGGTGATTTGGAAACGGGATTCGCCGTGCAAGAGT 1140
Db
1081 CGATCAGGCTCCCGCGCGGGAAACGGTGATTTGGAAACGGGATTCGCCGTGCAAGAGT 1140
QY
1141 GAGTAAGTACCGCCTATAGACTCTATATAGGACACACCCCTTTGGCTCTTATGCAATGCTATA 1200
Db
1141 GAGTAAGTACCGCCTATAGACTCTATATAGGACACACCCCTTTGGCTCTTATGCAATGCTATA 1200
QY
1201 CTGTTTTTGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGCTATAGCT 1260
Db
1201 CTGTTTTTGGCTTGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGCTATAGCT 1260
QY
1261 TAGCCTATAGGTGGGTTATTGACCAATATTGACCACTCCCCCTATTGGTGACGATACTT 1320
Db
1261 TAGCCTATAGGTGGGTTATTGACCAATATTGACCACTCCCCCTATTGGTGACGATACTT 1320
QY
1321 TCCATTACTAATCCATAAATGCTCTTTGCCCAAACTATCTCTATTGGCTATATGCCAA 1380
Db
1321 TCCATTACTAATCCATAAATGCTCTTTGCCCAAACTATCTCTATTGGCTATATGCCAA 1380
QY
1381 TACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTACAGGATGGGTCCATTTAT 1440
Db
1381 TACTCTGTCTTCAGAGACTGACACGACTCTGTATTTTACAGGATGGGTCCATTTAT 1440
QY
1441 TATTTACAAATTACATATACAAACAGCGCTCCCGCTGCCCGCAGTTTATTATAACA 1500
Db
1441 TATTTACAAATTACATATACAAACAGCGCTCCCGCTGCCCGCAGTTTATTATAACA 1500
QY
1501 TAGCGTGGGATCTCCGACATCTCGGATAGTGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
Db
1501 TAGCGTGGGATCTCCGACATCTCGGATAGTGTTCGGGACATGGGCTCTTCTCCGGTAGC 1560
QY
1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCAGCGGCTCATGGTCTCGGC 1620
Db
1561 GCGGAGCTTCCACATCCGAGCCCTGGTCCCATCCGTCAGCGGCTCATGGTCTCGGC 1620
QY
1621 AGCTCCTGTCTCTAACAGTGGAGGCGAGACTTAGGCACAGCAACAATGCCACCAACC 1680
Db
1621 AGCTCCTGTCTCTAACAGTGGAGGCGAGACTTAGGCACAGCAACAATGCCACCAACC 1680
QY
1681 AGTGTCCGCAAAAGCCGCTGGCGGTAGGATGTGTCTGAAATGAGCTCGGAGATTGG 1740
Db
1681 AGTGTCCGCAAAAGCCGCTGGCGGTAGGATGTGTCTGAAATGAGCTCGGAGATTGG 1740
QY
1741 GCTCGCACCTGGACGAGATGGAAGACTTAAGCAGCGGCAGAAAGATGACGGCAGCT 1800
|||||

Db
1741 GCTCGCACCTGGACGAGATGGAAGACTTAAGCAGCGGCAGAAAGATGACGGCAGCT 1800
QY
1801 GAGTTGTTGTATTCTGATAAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTTTAAACGGTGG 1860
Db
1801 GAGTTGTTGTATTCTGATAAGAGTCAAGGTAACTCCCGTTGCGGTGCTGTTTAAACGGTGG 1860
QY
1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACACAGACATAATAGCT 1920
Db
1861 AGGGCAGTGTAGTCTGAGCAGTACTCGTTGCTGCGCGCGCGCCACACAGACATAATAGCT 1920
QY
1921 GACAGACTAAACAGACTGTTCCCTTTCCATGGGTCTTTTCTGCACTCAACCGTCTGTCGACGAA 1980
Db
1921 GACAGACTAAACAGACTGTTCCCTTTCCATGGGTCTTTTCTGCACTCAACCGTCTGTCGACCTA 1980
QY
1981 TTTCAAGCAATCATGATGCAATGAAGAGAGGCTCTGCTGTGTGTGCTGTGTGTGAGAG 2040
Db
1981 AGNAATTCATGCA-----GTGGAA 1998
QY
2041 CAGTCTTTCGTTTCGCGCAGCGCTAGGGAACCCACGTCAACCGGGGAAGTGGCGGCCACA 2100
Db
1999 CTCACATGCTTCCACCAAACTCTGAGGATCCACAGAGTCAGGGGTCTGTATCTTCTCTGC 2058
QY
2101 CTGTGCTCGAATTTGTTAGCCTCTCTCGCACACGAGGCGCCAAAGCAGAAACGCTCCAGCTGATCA 2160
Db
2059 TGGTGCTC----- 2067
QY
2161 ACACCAACGGCAGTTGGCACCTCAATAGCACCGGCCCTGAACTGCAATGATGATGCTCAACA 2220
Db
2068 -----CAGTTTCAGGAACAGTAAGAACCTGCTCCGAATATTGCT----- 2105
QY
2221 CCGGCTGTTGGCAGGGCTTTTCTATCACCAACAAGTTCAACTCTTCAGGCTGTCTCTGAGA 2280
Db
2106 ----- 2105
QY
2281 GGCTAGCAGCTGCGACCCCTTACCGATTTTGAACAGGGCTGGGCCCTATCAGTTATG 2340
Db
2106 -----CTCACATCTCGTCAATCTCGCGAGGACTGGGACC----- 2141
QY
2341 CCAACGGAAGCGGCCCCGACAGCGCCCTACTGTGTGGCACTACCCCCCAAAACCTTGGG 2400
Db
2142 ----- 2141
QY
2401 GTATTGTCCCGGAAGAGTGTGTGTGCTCCGTATATTGCTTCACTCCAGCCCCGTGG 2460
Db
2142 ----- 2141
QY
2461 TGGTGGGAACGACCGACAGGTCGGGCGGCCCACTTACAGCTGGGGTGAAATGATACGG 2520
Db
2142 ----- 2141
QY
2521 ACGTCTTCTGCTTAAACAATACAGGCCACCGCTGGGCAATTTGGTTTCGGTTGTACCTGGA 2580
Db
2142 ----- 2141
QY
2581 TGAATCTCACTGGATTCCACAAAGTGTGCGGAGCGCTCTTGTGTCTATCGGAGGGCGG 2640
Db
2142 ----- 2141
QY
2641 GCAACAACACCTGCACTGCCCACTGATTGCTTCCGCAAGCATCCGAGCGCCACATACT 2700
Db
2142 ----- 2141
QY
2701 CTCGCTGGGCTCCGGTCCCTGGATCACACCGAGGTGCTGTGTGATACCGGTATAGGC 2760
Db
2142 ----- 2141
QY
2761 TTTGGCATTTATCTTGTACCACTCAACTACACCATATTTTAAATCAGGATGTACGTGGAG 2820
Db
2142 ----- 2141
QY
2821 GGGTCGAACACAGGCTGGAAGCTGCTCTGCAACTGGACGCGGGCGGAACGTTGCGATCTGG 2880
Db
2142 ----- 2141

2881 AAGATAGGACAGGTCCGAGATCGATATGAGAGAACATCATCAGGATTCCTTAGGACCCC 2940
2142 -----CTGTGACGAACATGGAGAACATCATCAGGATTCCTTAGGACCCC 2186
2941 TGCTCGTGTACAGCGGGGTTTTCTTGTGTGACAGAAATCCCTCAACAATCCGAGAGTC 3000
2187 TGCTCGTGTACAGCGGGGTTTTCTTGTGTGACAGAAATCCCTCAACAATCCGAGAGTC 2246
3001 TAGACTCGTGTGAGCTTCTCAATTTCTAGGGGATCTCCGCTGTGCTTTGGCCAAA 3060
2247 TAGACTCGTGTGAGCTTCTCAATTTCTAGGGGATCTCCGCTGTGCTTTGGCCAAA 2306
3061 ATTTCGAGTCCCAACCTCCAAATCACTCACCAACCTCCTCTCTCCAAATTTGCTCGGTT 3120
2307 ATTTCGAGTCCCAACCTCCAAATCACTCACCAACCTCCTCTCTCCAAATTTGCTCGGTT 2366
3121 ATCGCTGGATGTCTGCGCGGTTTTATCATATTTCTCTTCATCTCTGCTGCTATGCTCA 3180
2367 ATCGCTGGATGTCTGCGCGGTTTTATCATATTTCTCTTCATCTCTGCTGCTATGCTCA 2426
3181 TCTTCTTATTGGTTCTTCTCGATTATCAAGGTATGTTGCCGTTTTGCTCTTAATTCAG 3240
2427 TCTTCTTATTGGTTCTTCTCGATTATCAAGGTATGTTGCCGTTTTGCTCTTAATTCAG 2486
3241 GATCAACAAACAACAGTACGGGACCATGCAAAAACCTGCAGACTCTCTGCTCAAGGCAACT 3300
2487 GATCAACAAACAACAGTACGGGACCATGCAAAAACCTGCAGACTCTCTGCTCAAGGCAACT 2546
3301 CTATGTTTCCCTCANGTGTGTACAAAACCTACGGATGAAATTTGCACTGTATTCCTCA 3360
2547 CTATGTTTCCCTCANGTGTGTACAAAACCTACGGATGAAATTTGCACTGTATTCCTCA 2606
3361 TCCCATCGCTCGGCTTTTCGCAAAATACCTATGAGGAGTGGGCTCAGTCCGTTCTCTT 3420
2607 TCCCATCGCTCGGCTTTTCGCAAAATACCTATGAGGAGTGGGCTCAGTCCGTTCTCTT 2666
3421 GGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTCTGTAGGGCTTTCCGCCACTGTTTGGC 3480
2667 GGCTCAGTTTACTAGTGCCATTTGTTTCAGTGGTTCTGTAGGGCTTTCCGCCACTGTTTGGC 2726
3481 TTTTCAGCTATATGATGATGTGTATTTGGGGCCAAAGTCTGTACAGATCGTGAAGTCCCT 3540
2727 TTTTCAGCTATATGATGATGTGTATTTGGGGCCAAAGTCTGTACAGATCGTGAAGTCCCT 2786
3541 TTATACCGCTGTACCAATTTCTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
2787 TTATACCGCTGTACCAATTTCTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG 2846
3601 AGCAAGTCTAGAAAGCGCCCAAGATATCAAGGATCCACTACGCGTTAGAGTCTCGTGA 3660
2847 AGCAAGTCTAGAAAGCGCCCAAGATATCAAGGATCCACTACGCGTTAGAGTCTCGTGA 2906
3661 TCAGCTTCGACTGTGCTTCTAGTGTGACCAATCTGTTGTTTGGCCCTCCCGCTGCCT 3720
2907 TCAGCTTCGACTGTGCTTCTAGTGTGACCAATCTGTTGTTTGGCCCTCCCGCTGCCT 2966
3721 TCCTTGACCTGGAAGTGCCTCCCACTGCTCTTCTTAATAAATGAGGAAATTTGCA 3780
2967 TCCTTGACCTGGAAGTGCCTCCCACTGCTCTTCTTCTTAATAAATGAGGAAATTTGCA 3026
3781 TCCGATTTCTGATAGTGTGCTATTTCTTCTGGGGGTGGGTGGGGAGGACAGCAAG 3840
3027 TCCGATTTCTGATAGTGTGCTATTTCTTCTGGGGGTGGGTGGGGAGGACAGCAAG 3086
3841 GGGAGGATTTGGAAGACAATAGCAGCATGTCTGGGAGCTCTTCGCTTCTCGCTAC 3900
3087 GGGAGGATTTGGAAGACAATAGCAGCATGTCTGGGAGCTCTTCGCTTCTCGCTAC 3146
3901 TGACTCGCTCGCTCGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAGGCGGT 3960
3147 TGACTCGCTCGCTCGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAGGCGGT 3206

3961 AATACGGTTATCCACAGAAATCAGGGGATAAACGACAGAAAGAAACATGTGAGCAAAAGGCCA 4020
3207 AATACGGTTATCCACAGAAATCAGGGGATAAACGACAGAAAGAAACATGTGAGCAAAAGGCCA 3266
4021 GCNAAGGCCAGGAACCGTAAAAAGCCCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCC 4080
3267 GCNAAGGCCAGGAACCGTAAAAAGCCCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCC 3326
4081 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT 4140
3327 CCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT 3386
4141 ATAAAGATACGAGCGGTTTTCCCTCGGAAGTCCCTCGTGCCTCTCTGTTCCGACCT 4200
3387 ATAAAGATACGAGCGGTTTTCCCTCGGAAGTCCCTCGTGCCTCTCTGTTCCGACCT 3446
4201 GCGCTTACGGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATG 4260
3447 GCGCTTACGGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATG 3506
4261 CTCACGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCA 4320
3507 CTCACGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCA 3566
4321 CGAACCCCGCTTCAGCCCCGACCGCTCCGCTTATCCGGTAACTATCGTCTCGAGTCAA 4380
3567 CGAACCCCGCTTCAGCCCCGACCGCTTCGCTTATCCGGTAACTATCGTCTCGAGTCAA 3626
4381 CCGGTTAAGACAGGACTTATCGCACTGCGACGACGACCTGTGTAAACAGGATTAGCAGAC 4440
3627 CCGGTTAAGACAGGACTTATCGCACTGCGACGACGACCTGTGTAAACAGGATTAGCAGAC 3686
4441 GAGGTATGAGGCGGTCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAG 4500
3687 GAGGTATGAGGCGGTCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAG 3746
4501 AAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG 4560
3747 AAGGACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG 3806
4561 TAGCTCTTATCCGCAAAACAAACCAACCGCTGTAGCGGTGGTTTTTGTGTTGCAAGCA 4620
3807 TAGCTCTTATCCGCAAAACAAACCAACCGCTGTAGCGGTGGTTTTTGTGTTGCAAGCA 3866
4621 GCAGATTACGCGCAGAAAAAAGAGTCTCAAGAAAGATCTTTTGTCTTCTACGGGCTC 4680
3867 GCAGATTACGCGCAGAAAAAAGAGTCTCAAGAAAGATCTTTTGTCTTCTACGGGCTC 3926
4681 TGAAGTTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAATGAGATTATCAAAAAG 4740
3927 TGAAGTTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAATGAGATTATCAAAAAG 3986
4741 GATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 4800
3987 GATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 4046
4801 TGAGTAAAATCTGCTGACAGTTTACCAATGCTTTAATCAGTGAGCCACTTATCTCAGGAT 4860
4047 TGAGTAAAATCTGCTGACAGTTTACCAATGCTTTAATCAGTGAGCCACTTATCTCAGGAT 4106
4861 CTGCTTATTTCTGTTTCAATCCATAGTTCCTGACTCCCCGCTGCTGTAGATAACTACGATACG 4920
4107 CTGCTTATTTCTGTTTCAATCCATAGTTCCTGACTCCCCGCTGCTGTAGATAACTACGATACG 4166
4921 GAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGC 4980
4167 GAGGGCTTACCATCTGGCCCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGC 4226
4981 TCCAGATTTTATCAGCAATAAACAACGACGCGGAGCGGCGCAGAGAGTGTCTCTGC 5040
4227 TCCAGATTTTATCAGCAATAAACAACGACGCGGAGCGGCGCAGAGAGTGTCTCTGC 4286
5041 AACTTTTCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGGAAGCTAGAGTAAGTAGTTC 5100


```
Db 4287 AACCTTTATCCGCTCCATCAGTCTAATTAATTTGTCGGGAGCTAGAGTAGTATTC 4346
QY 5101 GCCAGTTAATAGTTTCGCAACGTTGTTGCCATGTCTACAGGCATCGTGGTGTACGCTC 5160
Db 4347 GCCAGTTAATAGTTTCGCAACGTTGTTGCCATGTCTACAGGCATCGTGGTGTACGCTC 4406
QY 5161 GTCGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATC 5220
Db 4407 GTCGTTTGGTATGGCTTTCATTCAGCTCCGGTTCCCAACGATCAAGCGAGTTACATGATC 4466
QY 5221 CCCCATGTTGTGC AAAAGCGTTAGCTCCCTTCGGTCTCCGATCGTTGTGAGAAGTAA 5280
Db 4467 CCCCATGTTGTG AAAAGCGTTAGCTCCCTTCGGTCTCCGATCGTTGTGAGAAGTAA 4526
QY 5281 GTTGGCCGCGAGTGTATCACTCATGTGTTATGGCAGCACTGCATAATTCCTTACTGTCTAT 5340
Db 4527 GTTGGCCGCGAGTGTATCACTCATGTGTTATGGCAGCACTGCATAATTCCTTACTGTCTAT 4586
QY 5341 GCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTCAATTCGTGAGAATA 5400
Db 4587 GCCATCCGTAAGATGCTTTTCTGTGATCGTGTAGTACTCAACCAAGTCAATTCGTGAGAATA 4646
QY 5401 GTGTATGCGCGGACCGAGTTGCTTTGCGCGGCTCAATACGGGATAATACCGCGCCACA 5460
Db 4647 GTGTATGCGCGGACCGAGTTGCTTTGCGCGGCTCAATACGGGATAATACCGCGCCACA 4706
QY 5461 TAGCAGAACTTTAAAAAGTGTCTCATCATTTGGA AAAACGTTCTCGGGCGGAAAACTCTCAAG 5520
Db 4707 TAGCAGAACTTTAAAAAGTGTCTCATCATTTGGA AAAACGTTCTCGGGCGGAAAACTCTCAAG 4766
QY 5521 GATCTTACCGCTGTTGAGATCCAGTTGCGATGTAAACCCACTCGTGTGACCCCAACTGATCTTC 5580
Db 4767 GATCTTACCGCTGTTGAGATCCAGTTGCGATGTAAACCCACTCGTGTGACCCCAACTGATCTTC 4826
QY 5581 AGCATCTTTTACTTTCACAGCGTTTCTGCGTGAGCA AAAACAGGAAGCAAAAATGCGCG 5640
Db 4827 AGCATCTTTTACTTTCACAGCGTTTCTGCGTGAGCA AAAACAGGAAGCAAAAATGCGCG 4886
QY 5641 AAAAAAGGGAATAAGGCGGACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATA 5700
Db 4887 AAAAAAGGGAATAAGGCGGACACGGAATGTTGAATACTCATACTCTCTCTTTTCAATA 4946
QY 5701 TTATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTTTA 5760
Db 4947 TTATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGATTTTA 5006
QY 5761 GAAAAATAACAAATAAGGCGGACATTTCCCGGAAAGTGCCACCTGACGCTCTA 5820
Db 5007 GAAAAATAACAAATAAGGCGGACATTTCCCGGAAAGTGCCACCTGACGCTCTA 5066
QY 5821 AGAAACCATTTATTCATGACATTAACCTATAAAAATAGGCGGTATCACGAGGCCCTTTTCG 5880
Db 5067 AGAAACCATTTATTCATGACATTAACCTATAAAAATAGGCGGTATCACGAGGCCCTTTTCG 5126
QY 5881 TC 5882
Db 5127 TC 5128
```

RESULT 4

```
US-09-721-480-1
; Sequence 1, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4276
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pCMVII
US-09-721-480-1
```

```
Query Match 45.2%; Score 2658.4; DB 4; Length 4276;
Best Local Similarity 72.7%; Pred. No. 0;
Matches 4275; Conservative 1; Mismatches 1; Indels 1606; Gaps 1;
```

```
QY 1 TCGCGCGTTTCGTTGATGACGGTGAAACCTCTGACACATGACAGCTCCCGGAGACGGTCA 60
Db 1 TCGCGCGTTTCGTTGATGACGGTGAAACCTCTGACACATGACAGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCGTCAGGCGCGGTGAGCGGGTG 120
Db 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCGTCAGGCGCGGTGAGCGGGTG 120
QY 121 TTGGCGGGTGTCTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGCTGAGAGTGC 180
Db 121 TTGGCGGGTGTCTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGCTGAGAGTGC 180
QY 181 ACCATATGAGCTTTTTCGAAAAGCTTAGGCCCTCCAAAAAGCCTCTCCTACTACTTCTGG 240
Db 181 ACCATATGAGCTTTTTCGAAAAGCTTAGGCCCTCCAAAAAGCCTCTCCTACTACTTCTGG 240
QY 241 AATAGCTCAGAGCGCGAGSGCGCTCGGCTCTGCATTAATAAAAAAATTAAGTCAGCCA 300
Db 241 AATAGCTCAGAGCGCGAGSGCGCTCGGCTCTGCATTAATAAAAAAATTAAGTCAGCCA 300
QY 301 TGGGCGGAGAAATGGGCGGAACCTGGGCGGGAGGAAATTAATGGCTATTGGCCATTGCAT 360
Db 301 TGGGCGGAGAAATGGGCGGAACCTGGGCGGGAGGAAATTAATGGCTATTGGCCATTGCAT 360
QY 361 AGTTGTATCTATATCATATATGATGATACATTTATATGGCTCATGTCCATATGACGCCCA 420
Db 361 AGTTGTATCTATATCATATATGATGATACATTTATATGGCTCATGTCCATATGACGCCCA 420
QY 421 TGTGACATTTGATTTTGTGACTAGTTTATTAATAGTAATCAATTAACGGGGTCATTAGTTCAAT 480
Db 421 TGTGACATTTGATTTTGTGACTAGTTTATTAATAGTAATCAATTAACGGGGTCATTAGTTCAAT 480
QY 481 AGCCCATATATGAGGTTCCGCGTTACATAAATTAACGGTAAATGGCCGCTGGCTGACCG 540
Db 481 AGCCCATATATGAGGTTCCGCGTTACATAAATTAACGGTAAATGGCCGCTGGCTGACCG 540
QY 541 CCCAAGCGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600
Db 541 CCCAAGCGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600
QY 601 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCCACTTGGCAGTA 660
Db 601 GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCCACTTGGCAGTA 660
QY 661 CATCAAGTGTATCATATGCAAGTCCGCGCTTACATAATGACGTCAATGACGTAAATGGCCC 720
Db 661 CATCAAGTGTATCATATGCAAGTCCGCGCTTACATAATGACGTCAATGACGTAAATGGCCC 720
QY 721 GCCTGGCATTTATGCCAGTACATGACCTTACCGGACTTTTCTTACTTGGCAGTACATCTAC 780
Db 721 GCCTGGCATTTATGCCAGTACATGACCTTACCGGACTTTTCTTACTTGGCAGTACATCTAC 780
QY 781 GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACACCAATGGCGGTGGA 840
Db 781 GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACACCAATGGCGGTGGA 840
QY 841 TAGCGGTTTGACTCAGCGGGAATTTCCAGTCTCCACCCCAATTCAGCTCAATGGGAGTTTG 900
Db 841 TAGCGGTTTGACTCAGCGGGAATTTCCAGTCTCCACCCCAATTCAGCTCAATGGGAGTTTG 900
```

901 TTTTGGCACCAGAAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 960
Db |||||
901 TTTTGGCACCAGAAATCAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 960
Qy |||||
961 CAAATGGCGGTAGGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020
Db |||||
961 CAAATGGCGGTAGGCGGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020
Qy |||||
1021 CGTCAGATCGCTTGGAGAGCCCATCAACGCTGTTTGGACCTCCATAGAAGACACCGGGAC 1080
Db |||||
1021 CGTCAGATCGCTTGGAGAGCCCATCAACGCTGTTTGGACCTCCATAGAAGACACCGGGAC 1080
Qy |||||
1081 CGATCAGGCTCCGCGCGCGGAAACGGTGCAATTTGGAAACGCGGATCCCGTGCAGAGT 1140
Db |||||
1081 CGATCAGGCTCCGCGCGCGGAAACGGTGCAATTTGGAAACGCGGATCCCGTGCAGAGT 1140
Qy |||||
1141 GAGTAAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCATGTATA 1200
Db |||||
1141 GAGTAAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCATGTATA 1200
Qy |||||
1201 CTGTTTTTGGCTTGGCGCTTATACACCCCGCTCTTATGCTATAGTGATGTATAGCT 1260
Db |||||
1201 CTGTTTTTGGCTTGGCGCTTATACACCCCGCTCTTATGCTATAGTGATGTATAGCT 1260
Qy |||||
1261 TAGCCTATAGGTGTGGTTATTGACCAATTATGACCACTCCCTATTTGGTGACGATACTT 1320
Db |||||
1261 TAGCCTATAGGTGTGGTTATTGACCAATTATGACCACTCCCTATTTGGTGACGATACTT 1320
Qy |||||
1321 TCCATTACTAATCCATAACATGCTCTTTGGCACAACATATCTCTATTTGGCTATATGCCAA 1380
Db |||||
1321 TCCATTACTAATCCATAACATGCTCTTTGGCACAACATATCTCTATTTGGCTATATGCCAA 1380
Qy |||||
1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTTAT 1440
Db |||||
1381 TACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTACAGGATGGGTCCATTTAT 1440
Qy |||||
1441 TATTTACAAATTCACATATACAAACACGGCGTCCCGTCCCGCAGTTTATTAAACA 1500
Db |||||
1441 TATTTACAAATTCACATATACAAACACGGCGTCCCGTCCCGCAGTTTATTAAACA 1500
Qy |||||
1501 TAGCGTGGGATCTCCGACATCTCGGTAGTGTTCGGACATGGGCTCTTCCGGTAGC 1560
Db |||||
1501 TAGCGTGGGATCTCCGACATCTCGGTAGTGTTCGGACATGGGCTCTTCCGGTAGC 1560
Qy |||||
1561 GCGGAGCTTCCACATCCGAGCCCTGTCCCATCCGTCCAGCGGCTCATGGTCCGTCGGC 1620
Db |||||
1561 GCGGAGCTTCCACATCCGAGCCCTGTCCCATCCGTCCAGCGGCTCATGGTCCGTCGGC 1620
Qy |||||
1621 AGCTCCTTCTCTAAACAGTGGAGGCCAGACTTTAGGCACAGCACAATGCCACCAACC 1680
Db |||||
1621 AGCTCCTTCTCTAAACAGTGGAGGCCAGACTTTAGGCACAGCACAATGCCACCAACC 1680
Qy |||||
1681 AGTGTCCGACACAGCCGCTGGCGGTAGGGTATGTCTGTAATAATGAGCTCGAGATTGG 1740
Db |||||
1681 AGTGTCCGACACAGCCGCTGGCGGTAGGGTATGTCTGTAATAATGAGCTCGAGATTGG 1740
Qy |||||
1741 GCTCGCACCTGGACGAGATGGAAGACTTAAGCAGCGGAGAGAGATGAGGCGAGCT 1800
Db |||||
1741 GCTCGCACCTGGACGAGATGGAAGACTTAAGCAGCGGAGAGAGATGAGGCGAGCT 1800
Qy |||||
1801 GAGTTGTTGTATTTCTGATTAAGAGTCAAGAGTAACTCCCGTTGGGTGCTTTAAACGGTGG 1860
Db |||||
1801 GAGTTGTTGTATTTCTGATTAAGAGTCAAGAGTAACTCCCGTTGGGTGCTTTAAACGGTGG 1860
Qy |||||
1861 AGGGAGTGTAGTCTGAGCAGTACTGTTGCTGCGCGCGGCGGCCACACAGATATAGCT 1920
Db |||||
1861 AGGGAGTGTAGTCTGAGCAGTACTGTTGCTGCGCGCGGCGGCCACACAGATATAGCT 1920
Qy |||||
1921 GACAGACTAACAGACTGTTCTTCCATGGGTCTTTTCTGCACTACCGTCTGTCGACGAA 1980
Db |||||
1921 GACAGACTAACAGACTGTTCTTCCATGGGTCTTTTCTGCACTACCGTCTGTCGAC --- 1977
Qy |||||
1981 TTCAAGCAATCATGGATGCAATGAAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAG 2040

1978 ----- 1977
2041 CAGTCTTCTGTTTCGCCCGAGGCTAGCGAAACCCACGTCACCGGGGGAAGTCCGGCCACA 2100
Db |||||
1978 ----- 1977
2101 CTGTGTCTGGAATTTGTAGCTCTCGCACAGGCGCAAGCAGAAAGCTCCAGCTGATCA 2160
Db |||||
1978 ----- 1977
2161 ACACCAACGGCAGTTGGCACCTCAATAGCAGCGGCCCTGAACCTGCAATGATAGCCTCAACA 2220
Db |||||
1978 ----- 1977
2221 CCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCTGAGA 2280
Db |||||
1978 ----- 1977
2281 GGCTAGCCAGCTGCCGAGCCCTTACCGATTTTGACCAAGGCTGGGGCCCTTATCAGTTATG 2340
Db |||||
1978 ----- 1977
2341 CCAACGGAAGCGGCCCGGACGAGCGCCCTACTGTGSGCACTACCCCCCAAAAACCTTGGC 2400
Db |||||
1978 ----- 1977
2401 GTATTGTGCCCGGAAGAGTGTGTGTGTCGGGTATATTGCTTCACTCCCGAGCCCGTGG 2460
Db |||||
1978 ----- 1977
2461 TGGTGGGAACGACCGAGAGTGGCGCGCCACCTTACAGTGGGGTGAAATGATACGG 2520
Db |||||
1978 ----- 1977
2521 AGCTCTTCTCTTAACAATAACAGGCCACCGCTGGGCAATTGGTTCTGTTGTAACCTGGA 2580
Db |||||
1978 ----- 1977
2581 TGAACCTCAACTGGATTACCAAAAGTGTGGAGCGGCTCTTGTGTCTATCGGAGGGCGG 2640
Db |||||
1978 ----- 1977
2641 GCAACAACACCTGCACCTGCCACCTGATTGCTTCGCAAGCATCGGAGCGCACATACT 2700
Db |||||
1978 ----- 1977
2701 CTCGGTGGGCTCCCGTCCCTGGATCACACAGGTGCTGTGTCGACTACCCGTATAGGC 2760
Db |||||
1978 ----- 1977
2761 TTTGGCATTTATCTTGTACCATCAACTACACCATATTTTAAATCAGGATGTACGTGGAG 2820
Db |||||
1978 ----- 1977
2821 GGGTCCGAACACAGGCTGGAAGTGTCTGCAACTGGACGCGGGCGGAACGTTGCGATCTGG 2880
Db |||||
1978 ----- 1977
2881 AAGATAGGACAGGTCCGAGATCGATATGGAGAAACATCATCAGGATTCCTTAGGACCCC 2940
Db |||||
1978 ----- 1977
2941 TGCTCGTGTACAGGGGGGTTTTTCTTGTTCACAAAGATCTTCACAATACCGCAGATC 3000
Db |||||
1978 ----- 1977
3001 TAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGATCTCCCGTGTGTCTTTGGCCAAA 3060
Db |||||
1978 ----- 1977
3061 ATTTCAGTCCCAACCTCCAATCATCAACAACTCTCTGTCTCTCCAATTTGTCTGGTT 3120

Db 1978 ----- 1977
Qy 3121 ATGCTGATGTCTGCGCGTTTATCATATTCCTTCATCCTGCTGCTATGCTCA 3180
Db 1978 ----- 1977
Qy 3181 TCCTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTTTGTCTCTAAATTCCAG 3240
Db 1978 ----- 1977
Qy 3241 GATCAACAACACAGTACGGGACCATGCAAAACCTGCACGACTCCTGCTCAAGGCAACT 3300
Db 1978 ----- 1977
Qy 3301 CTATGTTTCCCTCATGTTGCTGTACAAACCTACGGATGAAATTGCACCTGTATTTCCCA 3360
Db 1978 ----- 1977
Qy 3361 TCCCATGCTCGGGCTTTCGCAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTT 3420
Db 1978 ----- 1977
Qy 3421 GGCTCAGTTTACTAGTGCCATTGTTTTCAGTGGTTGCTAGGGCTTTCGCCCACTGTTTGGC 3480
Db 1978 ----- 1977
Qy 3481 TTTCAGCTATATGGATGATGTGTATTTGGGGCCAAAGTCTGTACAGCATCGTGAGTCCCT 3540
Db 1978 ----- 1977
Qy 3541 TTATACCGCTGTACCAATTTCTTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCG 3600
Db 1978 ----- 1977
Qy 3601 AGCAAGCTAGAAAGCGCGCAAGATATCAAGGATCCACTACGCGTTTAGAGTCTGCTCA 3660
Db 1995 AGCAAGCTAGAAAGCGCGCCAGATATCAAGGATCCACTACGCGTTTAGAGTCTGCTCA 2054
Qy 3661 TCAGCCTCGACTGTGCTTCTAGTTCGCCAGCATCTGTTTGTGCCCCCTCCCCCGTCCCT 3720
Db 2055 TCAGCCTCGACTGTGCTTCTAGTTCGCCAGCATCTGTTTGTGCCCCCTCCCCCGTCCCT 2114
Qy 3721 TCCTTGACCTGGAGGTGCCACTCCACTGCTTCTCCCTTAATAAGAGAAATTGCA 3780
Db 2115 TCCTTGACCTGGAGGTGCCACTCCACTGCTTCTCCCTTAATAAGAGAAATTGCA 2174
Qy 3781 TCGCATTGTCTGATAGGTGTCTATTCTATTCTGGGGGGTGGGTGGGCGAGACAGCAAG 3840
Db 2175 TCGCATTGTCTGATAGGTGTCTATTCTATTCTGGGGGGTGGGTGGGCGAGACAGCAAG 2234
Qy 3841 GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGAGCTCTTCCGCTTCTCGCTCAC 3900
Db 2235 GGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGAGCTCTTCCGCTTCTCGCTCAC 2294
Qy 3901 TGACTCGCTCGCTCGTTCGCTCGGCGAGCGGTATCAGTCACTCAAGAGCGGT 3960
Db 2295 TGACTCGCTCGCTCGTTCGCTCGGCGAGCGGTATCAGTCACTCAAGAGCGGT 2354
Qy 3961 AATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCCA 4020
Db 2355 AATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCCA 2414
Qy 4021 GCAAAAGGCAGAAACCGTAAAGCGCGTTGCTGGCGTTTTCATAGGCTCGGCC 4080
Db 2415 GCAAAAGGCAGAAACCGTAAAGCGCGTTGCTGGCGTTTTCATAGGCTCGGCC 2474
Qy 4081 CCTGACGAGCATCAGAAATCGAGCTCAGTACAGGTGGGAAACCCGACAGGACT 4140
Db 2475 CCTGACGAGCATCAGAAATCGAGCTCAGTACAGGTGGGAAACCCGACAGGACT 2534
Qy 4141 ATAAAGATACGAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCT 4200
Db 2535 ATAAAGATACGAGCGCTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCT 2594

Qy 4201 GCCGTTACCGATACCTGTCGCCCTTTCTCCCTTCGGGAAGGTGGCGCTTCTCAATG 4260
Db 2595 GCCGTTACCGATACCTGTCGCCCTTTCTCCCTTCGGGAAGGTGGCGCTTCTCAATG 2654
Qy 4261 CTCAGCGTGTAGGTATCTCAGTTCGCTGAGTTCGCTCCAAAGCTGGGCTGTGTGA 4320
Db 2655 CTCAGCGTGTAGGTATCTCAGTTCGCTGAGTTCGCTCCAAAGCTGGGCTGTGTGA 2714
Qy 4321 CGAAACCCCGCTTACGCGCGCTTACCGCTTATTCGGTAACTATCGTCTTTGAGTCCAA 4380
Db 2715 CGAAACCCCGCTTACGCGCGCTTACCGCTTATTCGGTAACTATCGTCTTTGAGTCCAA 2774
Qy 4381 CCCGTAAGACACGACTTATTCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGC 4440
Db 2775 CCCGTAAGACACGACTTATTCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGAGC 2834
Qy 4441 GAGGTATGTAGCGCGTGTACAGAGTTCCTTGAAGTGGTGGCCTTAACTACGGCTACACTAG 4500
Db 2835 GAGGTATGTAGCGCGTGTACAGAGTTCCTTGAAGTGGTGGCCTTAACTACGGCTACACTAG 2894
Qy 4501 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG 4560
Db 2895 AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG 2954
Qy 4561 TAGCTCTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGTGGTTTTTTTGTGCAAGCA 4620
Db 2955 TAGCTCTTGATCCGGCAAAACAAACCAACCGCTGCTGAGCGTGGTTTTTTTGTGCAAGCA 3014
Qy 4621 GCAGATTACGCCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCCTTTTACGGGGTC 4680
Db 3015 GCAGATTACGCCGAGAAAAAAGGATCTCAAGAAGATCCTTTGATCCTTTTACGGGGTC 3074
Qy 4681 TGACGCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCAAGAGATTATCAAAAAAG 4740
Db 3075 TGACGCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCAAGAGATTATCAAAAAAG 3134
Qy 4741 GATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATGATATA 4800
Db 3135 GATCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATGATATA 3194
Qy 4801 TCAGTAAACCTTCGCTGACAGTTACCAATGCTTAACTCAGTGAAGCACTTATCTCAGCGAT 4860
Db 3195 TGAGTAAACCTTCGCTGACAGTTACCAATGCTTAACTCAGTGAAGCACTTATCTCAGCGAT 3254
Qy 4861 CTGCTCTATTTCGTTTCAATCATAGTTGCCCTGACTCCCGCTCGTGTAGATAACTACGATACG 4920
Db 3255 CTGCTCTATTTCGTTTCAATCATAGTTGCCCTGACTCCCGCTCGTGTAGATAACTACGATACG 3314
Qy 4921 GGAGGCTTACCATCTGGCCCCAGTGTCTGCANTGATACCGGAGACCCAGCTCAGCGGC 4980
Db 3315 GGAGGCTTACCATCTGGCCCCAGTGTCTGCANTGATACCGGAGACCCAGCTCAGCGGC 3374
Qy 4981 TCCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTGCG 5040
Db 3375 TCCAGATTATCAGCAATAAACACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTGCG 3434
Qy 5041 AACTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCCGGAAGCTPAGAGTAAAGTATTC 5100
Db 3435 AACTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCCGGAAGCTPAGAGTAAAGTATTC 3494
Qy 5101 GCCAGTTAATAGTTTGGCGAAAGTTGTTGCCATTTGCTACAGGCATCGTGGTGTACGCTC 5160
Db 3495 GCCAGTTAATAGTTTGGCGAAAGTTGTTGCCATTTGCTACAGGCATCGTGGTGTACGCTC 3554
Qy 5161 GTCTGTTGATGTGGCTTCAATTCAGTCTCCGTTCCCAACCATCAAGCGGAGTTTACATGATC 5220
Db 3555 GTCTGTTGATGTGGCTTCAATTCAGTCTCCGTTCCCAACCATCAAGCGGAGTTTACATGATC 3614
Qy 5221 CCCCATTGTTGCAAAAAAGCGGTTAGTCTCTTCCGCTCTCCGATCGTTGTTCAGAAATGA 5280
Db 3615 CCCCATTGTTGCAAAAAAGCGGTTAGTCTCTTCCGCTCTCCGATCGTTGTTCAGAAATGA 3674

QY 1275 GGGTTATTGACCAATTATTGACCACTCCCTAT-----TGGTGAGGATACTTTCC 1323
Db 3887 GCCTTTCTCTCCACAGGTCACCTCCAGGTCAACTGACCTCGGTTCTATCGAATTGG 3946
QY 1324 ATTACTAAATCCATAACATGGCTCTTTGGCCACAACCTATCTCTATTGGCTATATGCCAATAC 1383
Db 3947 CCTTGGTGGCCAGCTGAGCTCAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4006
QY 1384 TCTGTCCTTACAGAGACTGACACGGACTCTGTATTTTTCACAGGATGGGTCCTATTATTAT 1443
Db 4007 NNN 4066
QY 1444 TTACAAATTCACATATACAAACGCGCTCCCGCTCCCGCAGTTTTTATTAAACATAG 1503
Db 4067 NNN 4126
QY 1504 CGTGGGATCTCCGACATCTCGGTACGTGTTCCGGACATGGGCTCTTCTCCGCTAGCGC 1563
Db 4127 NNN 4186
QY 1564 GGAGCTTCACATPCCGAGCCCTGGTCCCATCCGTCACAGCGGCTCATGGTGGCTCGGCAGC 1623
Db 4187 NNN 4246
QY 1624 TCCCTTCTCTAACAGTGGAGGCCAGCTTAGGCAAGCAATGCGCACCATCCAGCT 1683
Db 4247 NNN 4306
QY 1684 GTCCGCAACAAGCCGCTGGCGGTAGGTATGTCTGAATAATGAGCTCGAGAGATTGGGCT 1743
Db 4307 NNN 4366
QY 1744 CGCACCTGACGAGATGGAAGACTTAAGCAGCGCCAGAAAGATGACGAGCAGCTGAG 1803
Db 4367 ATGGCCCAACTTGTATTATGACCTTAATAATGGTTACAAATAAGCAATAGCATCAAA 4426
QY 1804 TTGTTGTATCTGATGAAGAGTACAGGTAACTCCGTTGGGTGCTGTTAAACGTTGGAGG 1863
Db 4427 TTTCACAAATAAAGCAATTTTTTCACTGCAATCTAGTTGGGTTTGTCCAAACTCATCA 4486
QY 1864 GCAGTGTAGTCTGAGCAGTACTCGTTGCTGCCGCGCGCCACAGACATAATAGCTGAC 1923
Db 4487 TGTATCTTATCATGCTCGATCGGGAATTAATTGGGCGCGGAGTCGACCTGCAGCCCA 4546
QY 1924 AGACTAACAGACTGTTCTTTCATCGGTCTTTTCTGCACTACCGTGTGACAGAAATTC 1983
Db 4547 AGCTTGGCAGCCAGAAATCCGCGCGGTGGTTTTTGGGGGTGCGGGGTGTTTGGCAGCCA 4606
QY 1984 AAGCAATCATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTGCTGTGTGGAGCAG 2043
Db 4607 CAGACGCCGCTGTTCTGTTGCGGCGAGTACATGCGGTCCATGCCCAGGCGCATCCAAAA 4666
QY 2044 TCTTCG---TTTCGCCCAGCGCTAGCGAAACCCACGTCACCGGGGGAAGTGC CGGCCACA 2100
Db 4667 CAGTGGGTCTGCTCTGCTCAGTCAGTCGTTGAGCTCAGCCCGCAGCAAGCCCAAGAAAT 4726
QY 2101 CTGTGTCTGATTTGTAGCTCTCTCGACACAGCGCCAGAGAGACGTCCTCAGCTGATCA 2160
Db 4727 AACCCCAACGAACCAATAAACCATTCCTCCATGGGGGACCCCGTCCCTTAAACCCAGCGGCC 4786
QY 2161 ACACCAACCGCAGTTGGCACCTCAATAGCACGCGCCCTGAACTGCAATGATAGCTCAACA 2220
Db 4787 GTGGCTATGGCGGGCTTGGCGCCCGCAGCTGTGGCTCGAGCCCTGGGCTCTACCCGAAAC 4846
QY 2221 CGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTTTCAACTCTTCAGGCTGTCTCTGAGA 2280
Db 4847 TTGGGGTTTGGGTTGGGGAAGGAAGA-----AACGGGCGGTATTGGCCCCAA 4896
QY 2281 GGCTAGCAGCTGCCGACCCCTTACCGATTTTGAACAGGGCTGGGCGCTTATCAGTTATG 2340
Db 4897 TGGGGTCTCGGTGGGGTATCGACAGAGTGCAGCCCTGGGACCGAAACCCCGCGCTTTATGA 4956

QY 2341 CCAACGAGAGCGGCCCGACAGCGCCCTCTGCTGGCACTACCCCCAAACCTTTGCG 2400
Db 4957 ACAACGAGCCCAACACCCCTGCTTTTATTTGTTTATTTGCGCTCATAGCGGGGT 5016
QY 2401 GTATTGTGTCGCCGAGAGTGTGTGTGTCGGGTATTTGCTTCACTCCAGCCCGGTGG 2460
Db 5017 TCTTTCGGTATTGCTCTCTCCGTGTTTCAGTTAGCTCCCGCATCTCCGATCCGGG 5076
QY 2461 TGGTGGAAACGACGACAGGTGGGGCGGCCACCTAAGCTGGGGTGAATGATACGG 2520
Db 5077 CGTGGGTTTCCACTATCGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCGGCTT 5136
QY 2521 AGTCTCTTCCTTAAATAACAGGCCACCGTGGGCAATTTGGTTGCGTTGTACCTGGA 2580
Db 5137 CTGGGGCCATTGTGTACGCCCGACAGTCCGGCTCCGGATCGGACGATTTGCGTCGAT 5196
QY 2581 TGAATCTCAACTGGATTACCAAAAGTGTGGGAGCGCTCTCTTGTGTATCGGAGGGCGG 2640
Db 5197 CGACCTCGGCCCAAGCTGCATCATCGAAATTTGCCGTCAACCAAGCTCTGTAGAGTTGG 5256
QY 2641 GCAACAC--ACCTGCACTGCCCCACTGATTTGTTCCGCAAGCATCCGGACGCCACATAC 2699
Db 5257 TCAAGACCAATCGGAGCATATACGCCCGGAGCGCGGCGATCCTGCAAGCTCCGGATGC 5316
QY 2700 TCTCGGTGGGCTCCGGTCCCTGATCACACAGGTGCTGCTGACTACCGGTATAGG 2759
Db 5317 CTCGCTCGAAGTAGGGCTGCTGCTCCATACAGCCAAACACGGCTCTCAGAAAGAG 5376
QY 2760 CTTTGGCATTTACTTTGTACCATCAACTACCAT-----ATTAAAAATCAGGA 2808
Db 5377 ATGTTGGGACCTCGTATTGGGAATCCCGGAACATCGCTCGCTCCAGTCAATGACCGCT 5436
QY 2809 TGTACTGGGAGGGTGCAGACACAGGCTGGAGAGCTGCTGCAACTGGACGCGGGCGAAC 2868
Db 5437 GTTATCGGCCCATTTGTCGTGAGACATTTGTTGGAGCCGAAATCCGCGTGCACGAGGTGC 5496
QY 2869 GTTGGCATCTGAAGATAGGACAGTCCGAGATCGATATGAGAACATCACATCAGGAT 2928
Db 5497 CGGACTTCGGGCGAGTCTCGGCCCAAGCATCAGCTCATCGAGAGCTTCGGGACGGAC 5556
QY 2929 TC-----CTAGGACCCCTGCTCGGTGTTACAGGCGGGTTTTTCTTGTGTGAAGAATC 2981
Db 5557 GCACTGACGGTGTGCTCCATCAGTTTTGCCAGTGATACATATGGGATCAGCAATCGCG 5616
QY 2982 CTCACAAATACCGCAGAGTCTAGACTCGTGTGAGCTTCTCTCAATTTTCTAGG--GGGA 3038
Db 5617 CATATGAATACGCCCATGTAGTATTGACCGATTTCTTGGCGTCCGAAATGGGCGCAAC 5676
QY 3039 TCTCCGCTGTGTTTGGCCAAAATTCGCACTCCCAACCTCCAAATCACCTCACCAACCTCC 3098
Db 5677 CGCTGCTGTGCTAAGATCGGCGGAGGATCGCATCCATGGCTCCGGGACGGCTGC 5736
QY 3099 TGTCTCTCAATTTGCTCGGTTATCGCTGGATGTGTCTCGGGGTTTTATCATATTTCTC 3158
Db 5737 AGAACAGCGGGCAGTTCCGTTTTCAGCAGGCTCTTGCAAGGTGACACCTGTGACGCGCG 5796
QY 3159 TTCACTCTGCTGATGCTCATCTTCTTATTTGGTTCTTCTGAGTATCAAGTATGTTG 3218
Db 5797 GAGTGAATAGGTGAGGCTCTCGCTGAATTTCCCAATGTCAAGCACTTCCGGAATCGGG 5856
QY 3219 CCGTTTGTCTTAAATCCAGGATCAACAAACAGTACGAGGACCATGCAAAACCTGC 3278
Db 5857 AGCGGGCCGATGCAAGTGCCGATAAAC-ATTAACGATCTTTGTAGAAACCATCGCGCA 5915
QY 3279 ACGACTCTCTGCTCAAGGCAACTCTATGTTTCCTCTCATGTTGCTGTATCAA-----AACC 3331
Db 5916 GCTATTATCCCGCAGGACATATCCAGCCCTCTCATCATCGAAGCTGAAAGCAGCATTC 5975
QY 3332 TACGATGGAATTTGACCTGTATTTCCCATCCATCGCTCGGCTTTCGCAAAATACCT 3391
Db 5976 TTCGCCCTCCGAGAGTGCATCAGGTGCGAGACGCTGTGCAACTTTTTCGATCAGAACTT 6035
QY 3392 ATGGGAGTGGGCTCAGTCCGTTTCTTGGGCTCAGTTTACTAGTGCCATTTTGTTCAGTG 3451

Db	6036	CTCGACAGACGTGCGCGTAGTTTCAGGCTTTTTCATATCTCAITGCCCCCGAGCAGGGA	6095
Qy	3452	GTTTCGTAGGCGTTTCCCCCACTGTTTGGCTTTTCAGGCTATATGATGATGTTGTTGGG	3511
Db	6096	TCGCGGCACGCTGTGTACGCTGTAAAGCGGGTCGTCAGAGGTGCGTCGGTGTTCAGG	6155
Qy	3512	GCCAAAGTCTGTACAGCATCGTAGT-----CCCTTTATACCGCTGTGTACCAATTTCTTTT	3567
Db	6156	CCACACGCGTCACTTAATATGGAAGTGACCTCGGACCGCGCCCGACTGCACT	6215
Qy	3568	GTCTCTGGGTATACATTTAAGAAATTCAGACT-----CGAGCAAGTCTAGAAAGGCGCG	3621
Db	6216	GCCTGTTTCGAATTCGCGCAATGACAAGACGCTGGCGGGTTCGTGTATCATATAGAACTAA	6275
Qy	3622	CAAGATATCAAGGATCCACTACGCTTAGAGCTCGCTGATCAGCTCGACTGTCCTTCT	3681
Db	6276	AGACATGCAATATATTTCTTCGGGGACACCCGACAGCAACCGAGCAACGGGCGACCG	6335
Qy	3682	AGTTG--CCAGCCATCTGTTGTTTGGCCCTCCGCCGTCCTTCCCTTGACCCCTGAAAGGTG	3739
Db	6336	GGATGAAGCAGGCGACGCTCTGGTCAITTTTCGGCGAGACCGCTTCGCTGAGCGCG	6395
Qy	3740	CCACTCCCATGCTCTTCTTAATAAATGAGGAAATGTCATCGCATTTGTCGATGAGT	3799
Db	6396	ACGATGATCGGCCCTGCTCGCTTGGGTATTCGGAATCTTGACGCCCTCGCTCAAGCCTTC	6455
Qy	3800	GTCAATTTCTTCTGGGGGTGGGTCGGGCGAGCAGCAGCAAGGGGAGGATGGGAGACA	3859
Db	6456	GTCACTGGTCCCGCCACCAAAAGTTTCGGCGAAGCAGCGCCATATTCGCGCGCATGGCG	6515
Qy	3860	ATAGCAGGATGCTGGGGAGCTCTTCGCTTCTCGCTCACTGACTCGCTCGCTCGGTC	3919
Db	6516	GCCGACGCGCTGGGCTACGCTTCTGCTGGCGTTCCGACGCGAGGCTGGATGGCTTCCCC	6575
Qy	3920	GTTTCGGCTGGCGAGCGGATCAGCTCACTCAAAAGCGGTAATA CGGTTATCCACAGAA	3979
Db	6576	ATTATGATTTCTTCGCTTCGCGGCATCGGGATGCGCGTTCAGGCGCATGCTGTCC	6635
Qy	3980	TCAGGGATACACGGAAGACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGACCGT	4039
Db	6636	AGCGAGTAGATACCAACCATCAGGACAGCTAAAAAGGCCAGCAAAAGGCCAGGACCGT	6695
Qy	4040	AAAAAGCCGCGTGTGCTGGCGTTTTTCCATAGAGCTCCGCCCTCCCTGACAGCATCAAAA	4099
Db	6696	AAAAAGCCGCGTGTGCGTTTTTCCATAGAGCTCCGCCCTCCCTGACGAGCATCAAAA	6755
Qy	4100	AATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACGAGCGTTT	4159
Db	6756	AATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACGAGCGTTT	6815
Qy	4160	CCCCCTGGAGCTCCCTCGCGCTCTCCTGTTCCGACCTCGCGCTTACCGGATACCTG	4219
Db	6816	CCCCCTGGAGCTCCCTCGCGCTCTCCTGTTCCGACCTCGCGCTTACCGGATACCTG	6875
Qy	4220	TCCGCTTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCACTGCTAGGTATCTC	4279
Db	6876	TCGCGCTTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCACTGCTAGGTATCTC	6935
Qy	4280	AGTTCCGCTGAGTCTGCTCCAGCTGGGCTGTGTGCAGAAACCCCGCTTACGCC	4339
Db	6936	AGTTCCGCTGAGTCTGCTCCAGCTGGGCTGTGTGCAGAAACCCCGCTTACGCC	6995
Qy	4340	GACCGCTGGCTTATCCGTAATCTATCTGTTAGTCCACCCGCTAGACAGCTTA	4399
Db	6996	GACCGCTGGCTTATCCGTAATCTATCTGTTAGTCCACCCGCTAGACAGCTTA	7055
Qy	4400	TCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGGAGGATGTAGCGGTGCT	4459
Db	7056	TCGCCACTGGCAGCAGCACTGGTAAACAGGATTAGCAGGAGGATGTAGCGGTGCT	7115
Qy	4460	ACAGAGTTCTTGAAGTGGTGGCTTAACCTAGGCTACACTAGAAAGGACAGTATTTGGTATC	4519

Db	7116	ACAGAGTTCTTGAAGTGGTGGCTTAACCTAGGCTACACTAGAAAGGACAGTATTTGGTATC	7175
Qy	4520	TGCGCTCTCGTGAAGCCAGTTACTTCGGAAGAGAGTTGGTAGCTCTTCATCCGCGAAA	4579
Db	7176	TGCGCTCTCTGTAAGCCAGTTACTTCGGAAGAGAGTTGGTAGCTCTTCATCCGCGAAA	7235
Qy	4580	CAAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTTACGCGCAGAAAA	4639
Db	7236	CAAACCAACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTTACGCGCAGAAAA	7295
Qy	4640	AAAGGATCTCAAGAAAGATCTTTTGTATCTTTCTACGGGCTCTGACGCTAGTGAACGAA	4699
Db	7296	AAAGGATCTCAAGAAAGATCTTTTGTATCTTTCTACGGGCTCTGACGCTAGTGAACGAA	7355
Qy	4700	AATCAAGTTAAGGATTTTGGTCATGAGATATCAAAAAGGATCTTCACTAGATCTTT	4759
Db	7356	AATCAAGTTAAGGATTTTGGTCATGAGATATCAAAAAGGATCTTCACTAGATCTTT	7415
Qy	4760	TTAAATTTAAAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC	4819
Db	7416	TTAAATTTAAAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC	7475
Qy	4820	AGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTATCC	4879
Db	7476	AGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTCTATTTCTGTTATCC	7535
Qy	4880	ATAGTTGCTGACTCCCGCTCGTGTAGATAAATCACTACGAGGAGGCTTACCATCTGGC	4939
Db	7536	ATAGTTGCTGACTCCCGCTCGTGTAGATAAATCACTACGAGGAGGCTTACCATCTGGC	7595
Qy	4940	CCAGTGTCTGCAATGATACCGGAGACCCACGCTCAACCGGCTCCAGATTTATCAGCAATA	4999
Db	7596	CCAGTGTCTGCAATGATACCGGAGACCCACGCTCAACCGGCTCCAGATTTATCAGCAATA	7655
Qy	5000	AACCGCAGCGGAAAGGCGCGAGCGCAGAGAGTGCTCTGCAACTTTATCCCGCTCATC	5059
Db	7656	AACCGCAGCGGAAAGGCGCGAGCGCAGAGAGTGCTCTGCAACTTTATCCCGCTCATC	7715
Qy	5060	CAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTAGTTCCGCACTTAATAGTTTGGCG	5119
Db	7716	CAGTCTATTAAATTTGTCGGGAAGCTAGAGTAAGTAGTTCCGCACTTAATAGTTTGGCG	7775
Qy	5120	AACGTTGTTGCCATTCGTACAGGCACTCGTGTGTCTACGCTCGTCTGTTGGTATGGCTTCA	5179
Db	7776	AACGTTGTTGCCATTCGTACAGGCACTCGTGTGTCTACGCTCGTCTGTTGGTATGGCTTCA	7835
Qy	5180	TTACGCTCCGTTTCCCAACGATCAAGGCGAGTTAATGATCCCGCATGTTGTGCAAAAAA	5239
Db	7836	TTACGCTCCGTTTCCCAACGATCAAGGCGAGTTAATGATCCCGCATGTTGTGCAAAAAA	7895
Qy	5240	GCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGTTGCGCCGAGTTTATCA	5299
Db	7896	GCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGTTGCGCCGAGTTTATCA	7955
Qy	5300	CTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCTGTCATGCCATCCGTAAGTCTTT	5359
Db	7956	CTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCTGTCATGCCATCCGTAAGTCTTT	8015
Qy	5360	TCGTGATGCTGGTGAATCTCAACCAAGTCAATCTCTGAGAAATAGTGTATGCGCGACCGAGT	5419
Db	8016	TCGTGATGCTGGTGAATCTCAACCAAGTCAATCTCTGAGAAATAGTGTATGCGCGACCGAGT	8075
Qy	5420	TGCTCTTCCCGCGCTCAACACGGGATAATCCGCGCCACATAGCAGAACTTTTAAAGTG	5479
Db	8076	TGCTCTTCCCGCGCTCAACACGGGATAATCCGCGCCACATAGCAGAACTTTTAAAGTG	8135
Qy	5480	CTCATCTTGGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA	5539
Db	8136	CTCATCTTGGAAAAAGCTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA	8195
Qy	5540	TCAGTTCGATGTAAACCCACTCGTGCAACCACTGATCTTACGATCTTTTACTTTTACC	5599
Db	8196	TCAGTTCGATGTAAACCCACTCGTGCAACCACTGATCTTACGATCTTTTACTTTTACC	8255

QY 5600 AGCGTTTCTGGGTGACCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCG 5659
DB 8256 AGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAGGGAATAAGGGCG 8315
QY 5660 ACACGGAAATGTTGAATACTCATACTCTTCTTTTCAATATATTATTGAAGCATTTATCAG 5719
DB 8316 ACACGGAAATGTTGAATACTCATACTCTTCTTTTCAATATATTATTGAAGCATTTATCAG 8375
QY 5720 GGTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAAACAAATAGGG 5779
DB 8376 GGTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAATAAAACAAATAGGG 8435
QY 5780 GTTCCGGCACAATTTCCCGAAAGTGCACCTGACGTCTAGAAACCAATATTATCATG 5839
DB 8436 GTTCCGGCACAATTTCCCGAAAGTGCACCTGACGTCTAGAAACCAATATTATCATG 8495
QY 5840 ACATTAACTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC 5882
DB 8496 ACATTAACTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC 8538

RESULT 6

US-09-299-141-1
; Sequence 1, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: SONG, SIHONG
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; EARLIER FILING DATE: 1998-04-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6565
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PLASMID C-AT
US-09-299-141-1

Query Match 38.4%; Score 2256.6; DB 3; Length 6565;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 3676; Conservative 0; Mismatches 1689; Indels 276; Gaps 13;

QY 495 GTTCCCGTTACATAACTTACGGTAAATGGCCGCTGGCTGACCGCCCAACGACCCCG 554
DB 176 GTTACCGTTACATACTTACGGTAAATGGCCGCTGGCTGACCGCCCAACGACCCCG 235
QY 555 CCCATTGACGTCAATATACGTATGTTCCCATAGTAAGCCCATAGGACATTTCCATTG 614
DB 236 CCCATTGACGTCAATATACGTATGTTCCCATAGTAAGCCCATAGGACATTTCCATTG 295
QY 615 ACCTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGATCA 674
DB 296 AGCTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGATCA 355
QY 675 TATGCCAAGTCGCCCCCTATTGACGTCAATGACGTAAATGCCCGCTGGCATTTATGC 734
DB 356 TATGCCAAGTCGCCCCCTATTGACGTCAATGACGTAAATGCCCGCTGGCATTTATGC 415
QY 735 CCAGTACATGACCTTACGGGACTTTTCTACTTGGCAGTACATCTAGCTATTAGTCATCG 794
DB 416 CCAGTACATGACCTTATGGGACTTTTCTACTTGGCAGTACATCTAGCTATTAGTCATCG 475
QY 795 TATTACATGGTGATCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC 854

DB 476 TATTACCATGGTGATCGGTTTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTC 535
QY 855 ACGGGATTTTCAAAGTCTCCACCCCATTCAGCGTCAATGGGAGTTTGTTTTGGCACCACAAA 914
DB 536 ACGGGATTTTCAAAGTCTCCACCCCATTCAGCGTCAATGGGAGTTTGTTTTGGCACCACAAA 595
QY 915 TCAAACGGGACTTTCCAAAATGCTGAATTAACCCCGCCCGTTGAGCGCAATAGGCGGCTAG 974
DB 596 TCAAACGGGACTTTCCAAAATGCTGAATTAACCCCGCCCGTTGAGCGCAATAGGCGGCTAG 655
QY 975 GCGGTACCGTGGGAGTCTATATAAGCAGAGCTCGTTTTAGTGAACCGTCAGATCCGCTG 1034
DB 656 GCGGTACCGTGGGAGTCTATATAAGCAGAGCTCGTTTTAGTGAACCGTCAGATCCGCTG 715
QY 1035 GAGAGCCCATCAACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGACCTCCG 1094
DB 716 GAGAGCCCATCAACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGACCTCCG 775
QY 1095 CGGCCGGGAACGGTGCAATTGGAACCGGATTCGCCGTGCCAAGAGTGAGTAAGTACCGC 1154
DB 776 GACTCTAGAACTAGTGGATCCCCCGGCTGCAGGAATTCGATATCAAGCTTGGGGATTTT 835
QY 1155 CTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATACTGTTTTTGGCTTG 1214
DB 836 CAGGCACACCACTGACCTGGGACAGTGAATCGACATGCCGCTCTTCTCTCGTGGGCG 895
QY 1215 GGGCTATACACCCCGCTCTTATGCTATAGGTATGGTATAGTTAGCTTATAGGTGT 1274
DB 896 ATCCTCTCTGCTGGCAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
QY 1275 GGGTTATTGACCACTATTGACCACTCCCCCTATTGTTGGAGCATACTTTTCATTACTAATCC 1334
DB 956 GGAGATGCTGCCAAGAACAGATATATCCACCATGATCAGATCACCACCACTTCAAC 1015
QY 1335 ATAAACATGCTCTTTGGCCACAACTATCTCTATTGGCTATATGCAATACTCTGTCCTTCA 1394
DB 1016 AGATCACCCCAACCTGGCTGAGTTGGCTTACGCTTATACCGCAGCTGGCAGACACCG 1075
QY 1395 GAGACTGACACGGACTCTGTTATTTTACAGGATGGGCTCCATTTATTTTACAA--ATT 1452
DB 1076 TCCAAACAGCACCAATATCTTCTCTCCCACTGAGCATCGCTACAGCTTTTGAATGCTC 1135
QY 1453 CACATATACAAACGCGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1512
DB 1136 TCCCTGGGACCAAGGCTGACACTCAGATGAAATCCTGGAGGCTGAAATTTCAACCTC 1195
QY 1513 TCCGACATCTCGGTAAGTGTTCGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCC 1572
DB 1196 ACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCCTCGTACCCCTCAAC 1255
QY 1573 ACATCGAGCCCTGGTCCCATCCGTCACGCGCTCATGTGCTGGGAGCTCCTTGCTC 1632
DB 1256 CAGCCAGACAGCAGCTCCAGCTGACCAACCGCAATGGGCTGTCTCAGCGAGGCGCTG 1315
QY 1633 CTAAACAGTGGAGCCAGACTTAGGCACAGCAATGCCCAACCAACCAAGTGTGCCGAC 1692
DB 1316 AAGCTAGTGAATGAAGTTTTTGGAGGATGTTAAAAGTTGTACCATCAGAGCCTTCACT 1375
QY 1693 AAGGCGCTGGCGTAGGGTATGTGTGTAATAAGCTCGAGATTTGGGCTCGACCTGG 1752
DB 1376 GTCMACTTCGGGACACCGAAGAGGCCAAGAACAGATCAACGATTCGTTGGAGAGGGT 1435
QY 1753 AGCAGATGGAAGACTTAAAGCAGCGGAGAGAGATCAGGCAGCTGAGTTGTGTAT 1812
DB 1436 ACTC--AAGGGAATAATTGTGATTTTGTCAAGGAGCTTGACAGAGACACAGTTTTTTGCTC 1493
QY 1813 TCTGATAAGAGTCAGAGGTAACTCCCGTTGCGTGTGTTAAACGGTGGAGGCGAGTGTAG 1872
DB 1494 TGGTGAATTAATCTTCTTTTAAAGGCAATGGG---AGAGACCTTTGAAGTCAAGGACA 1550
QY 1873 TCTGAGCAGTACTCGTTGCTGCGCGCGCCACCAAGACATATAGCTGACAGACTAACA 1932
DB 1551 CCGAGGAAGAGGACTTCCACGTGGACAGGTGACCAACCGTGAAGGTGCCTATGATGAAGC 1610

APPLICANT: Hauswirth, William W.
TITLE OF INVENTION: Humanized Green Fluorescent Protein
TITLE OF INVENTION: Humanized Green Fluorescent Protein
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT NUMBER: US/08/893,327
FILING DATE:
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/588,201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UFLA:062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 6280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:

NAME/KEY: CDS

LOCATION: 988..1728

US-08-893-327-19

Query Match

Best Local Similarity

Matches 3538; Conservative

0; Mismatches 1783; Indels 79; Gaps 16;

Score 2229.2; DB 3; Length 6280;

Pred. No. 0;

495 GTTCCCGGTACATACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 554

176 GGTACCCGTTACATACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 235

555 CCCATTGAGTCAATTAATGACGTATGTTCCCATAGTAAGCCCAATAGGGACTTTTCCATTG 614

236 CCCATTGAGTCAATTAATGACGTATGTTCCCATAGTAAGCCCAATAGGGACTTTTCCATTG 295

615 ACCTCAATGGGTGGAGTATTACGGTAAATGGCCCACTTGGCAGTACATCAAGTGTATCA 674

296 ACCTCAATGGGTGGAGTATTACGGTAAATGGCCCACTTGGCAGTACATCAAGTGTATCA 355

675 TATGCCAAGTCCGCCCTTATGACGTCAATGACGTAAATGGCCCGCTGGCATTATGC 734

356 TATGCCAAGTCCGCCCTTATGACGTCAATGACGTAAATGGCCCGCTGGCATTATGC 415

735 CCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTAGCTATTAGTATCGC 794

416 CCAGTACATGACCTTATGGGACTTTCTACTTGGCAGTACATCTAGCTATTAGTATCGC 475

795 TATTACCATGGTGTATGGGTTTTTGGCAGTACACCAATGGCGGTGATAGCGGTTTGACTC 854

476 TATTACCATGGTGTATGGGTTTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTC 535

855 ACGGGATTTCGAAGTCTCCACCCCATTTGACGTCAATGACGTGAGTTGTTTGGACCAAAA 914

536 ACGGGATTTCGAAGTCTCCACCCCATTTGACGTCAATGACGTGAGTTGTTTGGACCAAAA 595

915 TCAAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGTTCGACGCAAAATGGCGGTAG 974
596 TCAAACGGGACTTTTCCAAAATGTCGTAATAACCCCGCCCGTTCGACGCAAAATGGCGGTAG 655
975 GCGGTACCGGTGGAGGCTTATATAAGCAGAGAGTCTGTTAGTGAACCGTCAGATCGCTG 1034
656 GCGGTACCGGTGGAGGCTTATATAAGCAGAGAGTCTGTTAGTGAACCGTCAGATCGCTG 715
1035 GAGAGCCCATCCAGCGTGTGTTGACCTCATAGAAGACACCGGACCGATCCAGCGCTCCG 1094
716 GAGAGCCCATCCAGCGTGTGTTGACCTCATAGAAGACACCGGACCGATCCAGCGCTCCG 774
1095 CGGCGCGGAAACGGTGCATTGGAAACGGGATTTCCCGTGGCAAGAGTGAGTAAAGTACCGC 1154
775 -----GGACTCTAGAGGATCCGGTACTCGAGGAACCTGAAAAACCGAG 815
1155 CTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGATGCTATAGCTTAGCTTAGTGTG 1214
816 AAAGTTAACTGGTAAAGTTAGTCTTTTGTCTTTTATTTTTCAGGTCCCGATCCGGTGGT 875
1215 GGGCTTATACACCCCGCTCTTATGCTATAGTATAGTATAGCTTAGCTTAGCTTAGTGTG 1274
876 GTGCAAAATCAAGAACTGCTCTCAGTGATGTTGCTTTTCTTAGGCTCTAGCGAA 935
1275 GGGTTATTGACCACTTATTGACCACTCCCTTATTGGTGACGATATCTTTTCCATTACTAATCC 1334
936 GTGTTACTTCTGCTCTAAAAGCTGCGGAATTTGACCGCGCGCGCCACCATGTTGCC 995
1335 ATAAATGCTCTTTGCCCAACTATCTCTATTGGCTATATGCAATATCTCTGCTTCA 1394
996 CAAGAAGAGAGGAGGATGATGAGCAAGGGCGAGGAACCTGTTCACTGGCGTGTCCCAAT 1055
1395 GAGACTGACACGACTCTGTTATTTTACAGGATGGGTCCATTATTATTACAAATTC 1454
1056 TCTGTGGAACCTGGATGGCGATGTAATGGGCACAAATTTTCTGACGGAGAGGGTGA 1115
1455 CATATACAAACACCGCTCCCGCTGCGCGAGTCTTTTATTAAACATAGCGTGGGTCTC 1514
1116 AGGTGATGCCACATACGGAAGCTCACCTGAAATTCATCTGCACCACTGGAAGCTCCC 1175
1515 CGACATCTCGGTGATGTTTCCGACATGCGGTCTCTCTCGGTAGCGGGAGCTTCCAC 1574
1176 TGTGCGCATGGCAAC-----ACTGCTCACTACCTTCTCTATGCGGTGCTGCTTTTCCA 1231
1575 ATCCGAGCCCTGCTCCATCCGTCAGCGGCTCATGCTCGCTCGGAGCTCTTGTCTCT 1634
1232 GATACCCAGACATATGAAGCAGCATGACTTTTCAAGAGCGCCATGCCGAGGCTATG 1291
1635 AACAGTGGAGGCGCAGACTTTAGGCACAGCACAATGCCACACCACTGTTGCCGCAAA 1694
1292 TGCAGGAGAGAACCATCTTTTCAAAG-----ATGACGGGAACCTACAAGACCCGCGTGA 1347
1695 GCGGTGGCGGTAGGTATGTTCTGAAATAGCTCGGAGATGGGTGCGACCTGGAC 1754
1348 GTCAAGTTGCAAGGTGACACCCCTGTTGAATAGAATCGAGCTGAAGGGCATTTGACTTAAG 1407
1755 GCAGATGGAGACTTTAAGCAGCGGAGAGATGACGAGGAGCTGAGTTGTTGTTATTC 1814
1408 GAGATGGAAACATCTCGGCCCAAGCTGGAATACAACTTCAACTCCCAATTTGTATC 1467
1815 -TGATAAGAGTCAGAGGTAACCTCCCGTTCGGTGTCTTTAAACGGTGGAGGGCAGTGTAGT 1873
1468 ATCATGGCCGACAGCAAAAGAAATGCAATCAAGTCAACTTCAAGATCAGACACAAT 1527
1874 CTGAGCAGTATCTGTTGCTGGCGCGCGCCACACAGACATAATAGCTGACAGACTAACAG 1933
1528 GAGATGATCCGTGCGAGCTGGCGCCACCATTTATCAACAGAACACTCCAATTCGGCGAGC 1587
1934 ACTGTTCTCTTCCATGGGTCTTTCTGAGTACCTCGCTCGACGAAATTCACGCAATCAT 1993
1588 CTTGTCTCTCCACAGCAACCAATACCTGTTCCACCGAGTCTGCTCCCTGTCTAAAGATCCC 1647


```
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UFLA:062\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 988..1728
; US-08-893-327-17

Query Match      37.8%; Score 2225.6; DB 3; Length 6280;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 3518; Conservative 0; Mismatches 1809; Indels 67; Gaps 14;

QY 495 GTTCCGCGTTACATAACTTACGGTAATAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 554
Db 176 GGTACCCGTTACATAACTTACGGTAATAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 235
QY 555 CCCATTGACGTCNATAATGACGATGTTCCCATAGTAAGCCCAATAGGAGCTTTCCATTG 614
Db 236 CCCATTGACGTCNATAATGACGATGTTCCCATAGTAAGCCCAATAGGAGCTTTCCATTG 295
QY 615 AGTCNATGGGTGGAGTATTACGGTAATAATGGCCCACTTGGCAGTACATCAAGTGATCA 674
Db 296 AGTCNATGGGTGGAGTATTACGGTAATAATGGCCCACTTGGCAGTACATCAAGTGATCA 355
QY 675 TATGCCAAGTCCGCCCCCTATTGACGTCNATGACGGTAATAATGGCCCGCTGGCATTATGC 734
Db 356 TATGCCAAGTACGCCCCCTATTGACGTCNATGACGGTAATAATGGCCCGCTGGCATTATGC 415
QY 735 CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAGTATTAGTCAATCGC 794
Db 416 CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAGTATTAGTCAATCGC 475
QY 795 TATTACCATGGTGATCGGTTTTGGCAGTACACCAATGGCGTGGATAGCGTTTGACATC 854
Db 476 TATTACCATGGTGATCGGTTTTGGCAGTACATCAATGGCGTGGATAGCGTTTGACATC 535
QY 855 ACGGGGATTTCGAAGTCTCCACCCCATTCGACGTCAATGGGAGTTGTTTTGGCACCACAAA 914
Db 536 ACGGGGATTTCGAAGTCTCCACCCCATTCGACGTCAATGGGAGTTGTTTTGGCACCACAAA 595
QY 915 TCNACGGGACTTTCNAAAATGCTTAATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTG 974
Db 596 TCNACGGGACTTTCNAAAATGCTTAATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTG 655
QY 975 GGTGTACGCTGGGAGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTG 1034
Db 656 GGTGTACGCTGGGAGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTG 715
QY 1035 GAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCG 1094
Db 716 GAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGACCGATCCAGCCTCC- 774
QY 1095 CGGCCGGGAACGGTGATTTGGAAACGGGATTCCTCGTGCAGAGTACGTAAGTACCGC 1154
Db 775 -----GGACTAGAGGATTCGGTACTCGAGGAACCTGAAACACCG 815
QY 1155 CTATAGACTCTATAGGCACACCCCTTTGGCTCTATGTCATGCTATCTGTTTTGGCTTG 1214
Db 816 AAGTTTAAGTGAAGTTTGAAGCTTTTGTCTTTTATTTTTCAGTCCCGATCCGGTGGTG 875
QY 1215 GGGCCTTATACACCCCGCTCTTATGCTATAGTGTATAGTATAGCTTATAGCTTATAGGTT 1274
Db 876 GTGCAATCAAGAACTGCTCTCAGTGGATGTTGCTTTTACTTCTAGGCTGTACGGAA 935
```

```
QY 1275 GGGTTATTGACCACTTATTGACCACTCCCTATTGGTGAGCATACTTTCATTTACTAATCC 1334
Db 936 GTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGCGCGCGCCGCCACCATGTGTGC 995
QY 1335 ATAAATATGGCTCTTTGGCCACAACTATCTCTATTGGCTATATGCTCAATATCTGTCTTCA 1394
Db 996 CAAGAAGAGAGGAGGTGATGAGCAAGGGCGAGGAACGTGTTCACTGGCGTGTCCCAAT 1055
QY 1395 GAGACTGACACGGAATCTGTATTTTACAGGATGGGTCATTTATTTTCAAAATCA 1454
Db 1056 TCTCGTGGAACTGGATGGCATGTGAATGGGCACAAAATTTCTGTGAGCGGAGAGGTGA 1115
QY 1455 CATATACAAACACGCGCTCCCGTGGCGCAGTTTATTTAAACATACGCTGGGATCTC 1514
Db 1116 AGGTGATGCCACATACGGAAGCTCACCTGAAATTCATCTGCACACCTGGAAGCTCCC 1175
QY 1515 CGACATCTCGGATGCTGTTCCGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCCAC 1574
Db 1176 TGTGCCATGGCCAAACACTGGTCACTACCTGACCTATGGCG--TGCAGTGTCTTTCCAGA 1233
QY 1575 ATCCGAGCCCTGGTCCCATCCGTCAGCGGCTCATGTGTCGTCGCGGAGCTCCTTGTCTCT 1634
Db 1234 TACCCAGACCATATGAAGCAGCATGACTTTTTTCAAGAGCGCCATGCCGAGGCTATGTG 1293
QY 1635 AACAGTGGAGGCGGACTTAGGCACAGCACATGCCACCAACACAGTGTGCCGACAA 1694
Db 1294 CAGGAGAGAACCATCTTTTTTCAAGATGACGGGAACCTACAAGACCCGCGCTGAAGTCAAG 1353
QY 1695 GSCCGTGGCGGTAGGGTATGTCTGTAATAATGAGCTCGGAGATTTGGGCTCGCACCTGGAC 1754
Db 1354 TTCGAAGGTGACACCTTGGTGAATAGATCGAGCTGAGGGCATTTGACTTTTAAGGAGGAT 1413
QY 1755 GCAGATGGAAGACTTAAGGACGCGGAGAGAGATGAGGAGCTGAGTGTGTTGTTATTC 1814
Db 1414 GGAACCATTTCTCGGCCACAAAGCTGGAATACAACTATACTCCACAAATGTGTACATCATG 1473
QY 1815 TGATAGAGTACAGAGTAACCTCCGTTGCGGTGCTGTTAAACGTTGAGGCGGAGTGTAGTC 1874
Db 1474 GCCGACAAGCAAAAGAAATGGGCATCAAGGTCAACTTCAAGATCAGACACAACATTTGAGGAT 1533
QY 1875 TGACGAGTACTCGTTGCTGCCGCGCGCCACAGACATAATAGCTGACAGACTAACAGA 1934
Db 1534 GGATCGGTGACGCTGCCGACCATTTATCAACAGAACACTCCAATCGGCGACGCGCCTGTG 1593
QY 1935 CTGTTCTTTTCATGGGCTTTTCTGAGTGA--CGTGTGTCAGCAATTTCAAGCAATCAT 1993
Db 1594 CTCTCTCCAGACAACCAATTTACCTGTCCACCCAGTCTGCCCTGTCTAAAGATCCCAACGAA 1653
QY 1994 GGATCAATGAAGAGAGGCTCTGCTGTGCTGTGCTGTGCTGTGAGGAGTCTTCGTTTC 2053
Db 1654 AAGAGAGACCATGGTCTGCTGGAGTTTGTGACGCTGTGGGATCACAATGGCATG 1713
QY 2054 GCCACGCGCTAGCGAAAACCCACGTCAACGGGGGAAAGTGCAGGCGCACTGTGTGTGATTT 2113
Db 1714 GACGAGCTGTAAGTGAAGCGCGCGGGGATCCAGACATGATAAGATACATTTGATGAGT 1773
QY 2114 TGTAGCTCTCTCGCACGAGCGGCCAAGCAGAAAGCTGACGTGATCAACACCAACGCGAG 2173
Db 1774 TTGGCAAAACCAACAACTAGAAATGCAATGAAAAAATGCTTTTATTTGTGAAATTTGTGATG 1833
QY 2174 TTGGCACCTCAATAGC--ACGGCCCTGAACCTGAATGATAGCTCAACACCGCTGGTTGG 2232
Db 1834 CTATTGCTTTATTGTAACCATTTAAGCTGCAATTAACAGTTTAAACAACAATTTGCA 1893
QY 2233 CAGGGCTTTTCTATCACCACAAGTTCACTTTTCAAGGCTGTCTGAGGCTGAGCCAGCT 2292
Db 1894 TTTCATTTTATGTTTCAGGTTTCAAGGGGAGGTGTGGGAGGTTTTTTTAGTCGACTCGAGCA 1953
QY 2293 GCCGACCCCTTACCGATTTTGAACGAGGCTGGGGCCCTTATCAGTTATGCCAACGAGCG 2352
Db 1954 GTGTGTTTTGCAAGAGGAAGCAAAAGCC---TCTCCACCCAGGCGCTGGAATTTGTCCA 2010
QY 2353 GCCCGGACGAGCGCCCTACTGCTGGCACTACCCCCCAAAACCTTGGCGGTATTTGTCGCG 2412
```

2011 CCCAAGTCGAAGGCAAGTGTGGTGTTCGACAGGAAGCAAAAAGCTCTCCACCAGGCT 2070
2413 CGAAGAGTGTGTGGTCCGGTATATGCTTCACTCCAGCCCGGTGGTGGGAACGA 2472
2071 GGAA---TGTTTCCACCCTGAGCAACCCCGCCAGCGTCTGTCAATGGCGAAT 2127
2473 CCACAGGTGGGGGGCCCACTACAGCTGGGGTGAATAATGATAGGAGCTCTCGTCC 2532
2128 CGAACCGCAGATGCAGTCGGGGCGCGGTCCAGGTCCACTTCGCATATTAAGGTGA 2187
2533 TTAACAATACAGGCACCGCTGGGCAATGGTTGGTGTGTAAGTGAATCAACTG 2592
2188 CGCGTGTGGCTCGAACCAGCAGCAGCCCTGCAGCCAATATGGGATCGGCAATGAACA 2247
2593 GATTCA-CCAAAGTGTGGAGCGCTCTCTGTGTGTCATCGAGGGGGGCGACACACC 2651
2248 GATGGATGTGACGAGTTCCTCGGCGCTTGGGTGGAGGGCTATTCGGCTATGACTGG 2307
2652 CTGCACTGCCCACTGATTGCTTCCGCAAGCATCCGAGCCACATACTCTCGGTGGC 2711
2308 GCACAACAGACATCGCTGCTGATGCCCGGTGTTCCGGCTGTCAGCGCAGGGGCGC 2367
2712 TCCGTCCTCGATACACCCAGGTGCTGGTGCATACCCGTATAGGCTTTGGCAATAT 2771
2368 CCGGTTCTTTTGTCAAGACC---GACCTGTCCGGTGCCCTGAATGAACCTGCAGGACGAG 2424
2772 CTTGTACCATCAACTACACCATATTTAAATCAGGATGTACGTGGAGGGGTGCAACAC 2831
2425 GCAGCGGGCTATCGTGGGTGGCCAGACGGGGTTCCTTGGCGAGCTGTGCTCGACGTT 2484
2832 AGGCTGGAAGCTCCCTCAACTCGAGCGGGGGGCAACGTTGCGATCTGGAAGATAGGGAC 2891
2485 GTCACTGAAGCGGAGGAGTCTGCTGCTATTCGGCGAAGTGGCGGGCAGGATCTCCTG 2544
2892 AGTCCGAGATCGATATGGAGAACATCAGATAGGATTCCTAGGACCCCTGCTGTGTTA 2951
2545 TCATCTCACCTGCTCTGCGGAGAAATATCCATCATGCTGATGCAATGCGCGGCTG 2604
2952 CAGCGGGGTTTTCTTGTGTGCAAGNATCCTCACAATCCGAGAGTCTAGACTCGTG 3011
2605 CATAAG-----CTTGATCCGGCTACCTGCCCCATTCGACCAACAGCGAAACATCGCAT 2657
3012 TGACCTTCTCTCAATTTTCTAGGGGATCTCCCGTGTGCTTGGCCAAAATTCGACATCC 3071
2658 CGAGCGACGATCTCGATGGAAGCCGCTCTTGTGATCAGATGATCTGACGAGA 2717
3072 CCAACCTCAATCACTACCAACCTCTGCTCTCAATTTGCTGCTGATTCGCTGATG 3131
2718 GCATCAGGGGCTCGCGCAGCCGAACTGTTGCGCAGGCTCAAGGCGCGATGCGCGCG 2777
3132 TGTCTGGGGTTTTATCATATTTCTTCACTATCTGCTGCTATGCTCTTCTTATG 3191
2778 CGAGGATCTGCTGACCCATGCGATGCTGCTGCTGCGCAATATCATGGTGAATAATG 2837
3192 GTTCTTTCTGATATCAAGATATGTTGCCGTTTGTCTTAATTCAGGATCAACAACA 3251
2838 CCGCTTTCTGGATTATCATGACTGTGGCGGCTGGGTGGCGGACGCTATCAGGACAT 2897
3252 ACCAGTACGGGACCATGCAAAAACCTGACGACTCTGCTCAAGGCAACTCTATGTTTCCC 3311
2898 AGGTTTGGGTACCCGCTGATATCTGAAAGAGCTTGGCGGGAATGGGCTGACCGCTTCC- 2956
3312 TCATGTTGCTGTAACAAACCTACGGATGGAATTCGACCTGATTTCCATCCCATCGTCC 3371
2957 ---TCGTGCTTACGGTATCGCGCTCCCGATTTCGCA-----GGCGATCGCTTCTAT 3006
3372 TGGGCTTTTCGAAAATACCTATCGGAGTGGGCTC-AGTCCGTTTCTTGGCTCAGTTT 3430
3007 CGGCTTCTTGACGAGTTCTTCTGAGGGGATCCGTGCACTAGAGCTCGCTGATCAGCCTCG 3066
3431 ACTAGTGCCATTTGTCAGTGGTTCGTAGGGCTTCCCGGCTGTTTGGCTTTTCACTAT 3490

3067 ACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGTGCCTTCTTGAAC 3126
3491 ATGGATGATGTGATTTGGGGCCAAAGTCTGTACAGCATCGTGAGTCCCTTTTATACGCT 3550
3127 CTGGAAGGTGCCACTCCCACTGCTTCTTAATAAATGAGGAATTCGATCGCATTTG 3186
3551 GTTACCAATTTTCTTTTGTCTCTGGGTATACATTTAAG--AATTACAGACTCGAGCAAGTC 3608
3187 CTGAGTAGTGTGCTATTCTATTCTGCGGGTGGGTGGGCGAGCAGCAAGAGGGAGGAT 3246
3609 TAGAAGGCGCGCAAGATATCAAGATCCACTACGCTTAGAGCTCGCTGATCAGCCTC 3668
3247 TGGGAAGACAATAGCAGGCACTGCTGGGAGAGATCTAGGAACCCCTAGTGATGAGTTGG 3306
3669 GACTGTGCTTCTAGTTGCCAGCATCTGTTGTTTGGCCCTCCCGTGCCTTCTTGTAC 3728
3307 CCACTCCCTCTCTGCGGCTGCTGCTCCTACTGAGGCGCGCGGCAAGCCCGGGGCTC 3366
3729 CTTGGAAGGTGCCACTCCCACTGCTTCTTCTTAATAAATGAGGAATTTGCACTGCAATTG 3788
3367 GGGCGACCTTTGGTCCCGCGGCTC-----AGTGAGCGAGCGAGCGCGCAGAG 3414
3789 TCTGAGTAGTGTGCTATTCTATTCTGGGGGTGGGTGGGCGAGGACAGCAAGGGGAGGA 3848
3415 AGGGAGTGCCAAACCCCGGCTGAGCCCTGCACTCACTCAAAAGGCGTAAATCGGCAAC 3474
3849 TTGGGAAGACAATAGCAGGCACTGCTGGGAGCTTCTCGGCTTCTCGCTCACTGACTCGC 3908
3475 GCGCGGGAGAGCGGTTTGGGTATTTGGGCGCTTCTCGGCTTCTCGCTCACTGACTCGC 3534
3909 TCGCTCGGTGCTTGGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATCGGT 3968
3535 TCGCTCGGTGCTTGGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGTAAATCGGT 3594
3969 TATCCACAGAAATCAGGGGATACCGAGGAAAGACATGTGAGCAAAAGGCGCAAAAGG 4028
3595 TATCCACAGAAATCAGGGGATACCGAGGAAAGACATGTGAGCAAAAGGCGCAAAAGG 3654
4029 CAGGAACCGTAAAGGCGGCTTCTGCGGCTTCTCATAGGCTCCGCGCCCTTGACG 4088
3655 CAGGAAACCGTAAAGGCGGCTTCTGCGGCTTCTCATAGGCTCCGCGCCCTTGACG 3714
4089 AGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAGAT 4148
3715 AGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAGAT 3774
4149 ACAGCGGTTTCCCTCGAAGCTCCTCGTGGCTCTCTGTTTCGACCCCTGCGCTTA 4208
3775 ACCAGCGTTTCCCGTGGAACTCCTCGTGGCTCTCTGTTTCGACCCCTGCGCTTA 3834
4209 CCGGATACCTGTCGCGCTTCTCCCTTCCGGGAAGCGTGGCGCTTCTCAATGCTCAGCT 4268
3835 CCGGATACCTGTCGCGCTTCTCCCTTCCGGGAAGCGTGGCGCTTCTCAATGCTCAGCT 3894
4269 GTAGGTATCTCAGTTCCGTTAGTGTGCTTCAAGCTGGGCTGTGTGACGAAACCC 4328
3895 GTAGGTATCTCAGTTCCGTTAGTGTGCTTCAAGCTGGGCTGTGTGACGAAACCC 3954
4329 CCGTTACGCGCGCTGCGCTTATCGGTAACTATGCTTGTAGTCCAAACCCGTTAA 4388
3955 CCGTTACGCGCGCTGCGCTTATCGGTAACTATGCTTGTAGTCCAAACCCGTTAA 4014
4389 GACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGCGAGGTATG 4448
4015 GACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGCGAGGTATG 4074
4449 TAGGCGGTCTACAGATTCTTTGAAGTGTGGCTTAACCTACGCTTACACTAGAAAGCAG 4508
4075 TAGGCGGTCTACAGATTCTTTGAAGTGTGGCTTAACCTACGCTTACACTAGAAAGCAG 4134
4509 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCGAAAGAGTTGGTAGTCTTT 4568
4135 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCGAAAGAGTTGGTAGTCTTT 4194

QY 4569 GATCCGGGAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCCAAGCAGCAGATTA 4628
DB 4195 GATCCGGGAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCCAAGCAGCAGATTA 4254
QY 4629 CCGCGAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGGTCTGAGCGTC 4688
DB 4255 CCGCGAGAAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGGTCTGAGCGTC 4314
QY 4689 AGTGGAAAGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCA 4748
DB 4315 AGTGGAAAGAAACCTCAGCTTAAGGATTTTGGTCATGAGATTTATCAAAAAGGATCTTCA 4374
QY 4749 CCTAGATCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATGAGTAAA 4808
DB 4375 CCTAGATCTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATGAGTAAA 4434
QY 4809 CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTAT 4868
DB 4435 CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTAT 4494
QY 4869 TTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGATAACTACGATACGGGAGGCT 4928
DB 4495 TTCGTTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGATAACTACGATACGGGAGGCT 4554
QY 4929 TACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCAGGCTCCAGATT 4988
DB 4555 TACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCACCAGGCTCCAGATT 4614
QY 4989 TATCAGCAATAAACAGCAGCGGAGCGGAGCGGAGAGAGTGTCTGCAACTTTAT 5048
DB 4615 TATCAGCAATAAACAGCAGCGGAGCGGAGCGGAGAGTGTCTGCAACTTTAT 4674
QY 5049 CCGCTCTCCATCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTAGTTGCGCAGTTA 5108
DB 4675 CCGCTCTCCATCAGTCTATTAATTTGTCGGGGAAGCTAGAGTAAGTAGTTGCGCAGTTA 4734
QY 5109 ATAGTTTGGCAACGTTGTTGCCATGTGCTACAGGCATCTGTTGTACCGCTCGTCTGTTTG 5168
DB 4735 ATAGTTTGGCAACGTTGTTGCCATGTGCTACAGGCATCTGTTGTACCGCTCGTCTGTTTG 4794
QY 5169 GTATGCTTTCATTTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATATCCCCATGT 5228
DB 4795 GTATGCTTTCATTTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATATCCCCATGT 4854
QY 5229 TGTGCAAAAAGCGGTAGTCTCTTCGTCCTCCGATCGTTGTGCAAGTAAGTTGGCGC 5288
DB 4855 TGTGCAAAAAGCGGTAGTCTCTTCGTCCTCCGATCGTTGTGCAAGTAAGTTGGCGC 4914
QY 5289 CAGTGTATCACTCATGTTATGGCAGCACTGCATTAATTTCTTACTGTCATGCCATCCG 5348
DB 4915 CAGTGTATCACTCATGTTATGGCAGCACTGCATTAATTTCTTACTGTCATGCCATCCG 4974
QY 5349 TAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCATTCTCAGAAATAGTGTATGC 5408
DB 4975 TAAGATGCTTTTCTGTGACTGTTGAGTACTCAACCAAGTCATTCTCAGAAATAGTGTATGC 5034
QY 5409 GCGCAGCCGAGTTGCTCTTCCCGCGGTCAATACGGGATAATACCGCGCCACATAGCAGAA 5468
DB 5035 GCGCAGCCGAGTTGCTCTTCCCGCGGTCAATACGGGATAATACCGCGCCACATAGCAGAA 5094
QY 5469 CTTTAAAGTGTCTCATCTATGGAACGTTCTTCGGGGGGAACCTCTCAAGATCTTAC 5528
DB 5095 CTTTAAAGTGTCTCATCTATGGAACGTTCTTCGGGGGGAACCTCTCAAGATCTTAC 5154
QY 5529 CGCTGTTGAGATCCAGTTCGATGTACCCACTCGTCACCCAACTGATCTTCAGATCTT 5588
DB 5155 CGCTGTTGAGATCCAGTTCGATGTACCCACTCGTCGACCCCACTGATCTTCAGATCTT 5214
QY 5589 TTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGG 5648
DB 5215 TTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGG 5274

QY 5649 GAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTATTGAA 5708
DB 5275 GAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCTCTTTTCAATATTATTGAA 5334
QY 5709 GCATTATATCAGGGTTATGTTCTCATGAGCGGATACATATTTTGAATGTTATTAGAAAAATA 5768
DB 5335 GCATTATATCAGGGTTATGTTCTCATGAGCGGATACATATTTTGAATGTTATTAGAAAAATA 5394
QY 5769 AACAAATAGGGTTCGCGCACATTTTCCCGCAAAAGTGCACCTGACGTCTTAAGAAACCA 5828
DB 5395 AACAAATAGGGTTCGCGCACATTTTCCCGCAAAAGTGCACCTGACGTCTTAAGAAACCA 5454
QY 5829 TTATTATCATGACATTAACCTATAAAATAGCGGTATCAGGAGCGCTTTCGTC 5882
DB 5455 TTATTATCATGACATTAACCTATAAAATAGCGGTATCAGGAGCGCTTTCGTC 5508

RESULT 9
US-08-893-327-15
; Sequence 15, Application US/08893327
; Patent No. 6020192
; GENERAL INFORMATION:
; APPLICANT: Zolotukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; TITLE OF INVENTION: Genes and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,327
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/588,201
; FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UFLA:062\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 988..1701
; US-08-893-327-15

Query Match 37.8%; Score 2220.6; DB 3; Length 6253;
Best Local Similarity 65.9%; Pred. No. 0;
Matches 3566; Conservative 0; Mismatches 1719; Indels 124; Gaps 19;

QY 495 GTTCCGCGTTACATACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 554
DB 176 GGTACCCGCGTTACATACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCG 235

Db 2345 TTCTTTTGTCAAG - ACCGACCTGTCCGGTCCCTGAATCAACTGTGAGGACGAGCG 2403
QY 2757 AGCTTTGGCAATTATCTGTGTAACAATACTACACCATATTTAAATCAGGATGTAAGTG 2816
Db 2404 CGCTATCTGTGGTGGCCACGACGGCGTTCCTTGGCGAGCTGTGCTGACGTTGTCACT 2463
QY 2817 GGAGGGTTCGAAACACAGGCTGGAAGCTGCTGCAATGAGACGGGGCGAAGTTCGGAT 2876
Db 2464 GAAGCGGAAGGACTGTGCTGCTATTGGGGCAAGTGC CGGGGAGGATCTCTGTCACT 2523
QY 2877 CTGGAAGATAGGACAGGTCGAGATCGATATGGAACAATACATCAGGATTCCTAGGA 2936
Db 2524 CACCTTGTCTCCCGAGAAAGTATCCATCATG - GCTGATGCAATGCGGGCTGCATA 2581
QY 2937 CCCTGCTGTGTTACAGGGGGTTTTTCTGTTGACAGAAATCTCAATACCGAG 2996
Db 2582 CGCTTGATCCGGCTACCTGCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCAC 2641
QY 2997 AGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGATCTCCGTTGTCTTGGC 3056
Db 2642 GTRACTCGAATGGAGCGGTCT - - - - - TGTGATCAGGATGATCTGGAC 2685
QY 3057 CAAAATTCGAGTCCCAACCTCCAATCACTCAACCACTCTGTCTCTCAATTTGTCT 3116
Db 2686 GAAGACATCAGGGC - - - - - TCGCGCCAGCGAACTGTTTCGCCAGGCTCAAGC 2735
QY 3117 GGTATTCGCTGATGTGTGCGGGCGTTTATCATATCTCTTCAATCTCTGCTGCTATGC 3176
Db 2736 GGCATGCGCGACGGCGAGGATCTGCTGCTGACCATGCGATGCTGCTTGCAGATAT 2795
QY 3177 CTCATCTTCTATTGTTCTCTCGGATATCAAGGTATGTGCGCCGTTTGTCTCTAAT 3236
Db 2796 CATGTGGAAATGGCGCTTTTCTCGATTCTGACTGTGGCGGCTGGGTGCGGA 2855
QY 3237 CCAGGATCAACAAACCAAGTACGGACCATGCAAAACCTGCAACACTCTGTCTAAGC 3296
Db 2856 CCGCTATCAGGACATAGCGTTGCTACCTCGTGATATTGCTGAAGACTTGGCGCGAATG 2915
QY 3297 AACTATGTTTCCCTCATGTTGCTGTACAAACCTACGATGGAAATTTGACCTGTATT 3356
Db 2916 GGCTGACCGGTTCC - - - - - TGTGCTTTTACGGTATCGCGCTCCCGATTTCGCA - - - - - G 2964
QY 3357 CCATCCCATCGCTCGGGCTTTTCGAAATACTATGGAGTGGGCTTC - AGTCCGTTT 3415
Db 2965 CGATCGCTTCTATCGCTTCTTGAGAGTCTTCTGAGGGATCCGTGACTAGACT 3024
QY 3416 CTCCTGGCTCAGTTACTAGTGCAATTTGTTCAAGTGTTCGTAGGGCTTTCCGCCACTGT 3475
Db 3025 CGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTCCTCCCTCC 3084
QY 3476 TTGGCTTTCAGCTATATGATGATGTGTTATGGGGGCCAGTCTGTACAGCATCTGAG 3535
Db 3085 GTGCTTCTCTGACCTCGAAGGTGGCACTCCCACTGTCTTCTCTAATAAATGAGAA 3144
QY 3536 TCCCTTTATACCGCTGTACCAATTTTCTTGTCTCTGGGTATACATTTAAG - AATTTC 3593
Db 3145 ATTGCAATCGATTGCTGATGAGTGTCAATTTCTGGGGGTGGGTGGGCGAGGAC 3204
QY 3594 AGACTCGAGCAAGTCTAGAAAGCGCGCAAGATATCAAGGATCCACTACGGTTAGAGC 3653
Db 3205 AGCAAGGGGAGGATTTGGGAAGACAATAGCAGGCATGCTGGGAGAGATCTAGGAACCCC 3264
QY 3654 TCGCTGATCAGCTCGACTGTGCTTCTAGTTGCGAGGCATCTGTTGTTGGCTCCCTCC 3713
Db 3265 TAGTGTAGGATTTGGCACTCTCCCTCTCTGCGCTCGCTCACTGAGGCGCGCCGGG 3324
QY 3714 CGTGCCTTCTTCACTCGAAGGTGCCACTCCCACTGTCTTCTCTAATAAATGAGGA 3773
Db 3325 CAAGCCCGGGCTCGGGGACCTTTGTTGCTGCCCGGCTC - - - - - AGTAGCG 3372
QY 3774 AATTGATCGCATGTGCTGATGAGTGTCAATTTCTGGGGGTGGGTGGGCGAGGA 3833
Db 3373 AGCGAGCGCGAGAGAGGAGTGGCCAAACCCCGCCCCCCCCCTGCGAGCCCTGCATT 3432

QY 3834 CAGCAAGGGGAGGATTTGGAAAGACAATAGCAGGCAATGCTGGGAGCTCTTCGGCTTCCT 3893
Db 3433 AATGAATTCGGCAACCGCGGGGAGAGCGGTTTGCATATTGGGCGCTCTTCGGCTTCCT 3492
QY 3894 CGCTCACTGACTCGCTCGCTCGCTCGGCTGCGCTGCGGAGCGGTATCAGCTCACTCAA 3953
Db 3493 CGCTCACTGACTCGCTCGCTCGCTCGGCTGCGGAGCGGTATCAGCTCACTCAA 3552
QY 3954 AGCGGTAATACGGTTTATCCAGAAATCAGGGGATAACGACAGAAAGAAATGTGAGCAA 4013
Db 3553 AGCGGTAATACGGTTTATCCAGAAATCAGGGGATAACGACAGAAAGAAATGTGAGCAA 3612
QY 4014 AAGCCAGCAAAAGGCGAGGAAACCGTAAAGAGCGCGCTGTGCTGGGTTTTTCCATAGC 4073
Db 3613 AAGGCCAGCAAAAGGCGAGGAAACCGTAAAGAGCGCGGTGTGCGGTTTTTCCATAGC 3672
QY 4074 TCGCGCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 4133
Db 3673 TCGCGCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 3732
QY 4134 CAGGACTATAAAGATAACAGGCGTTTCCCTTGGAAAGCTCCCTCGTGCGCTCTCTGTTC 4193
Db 3733 CAGGACTATAAAGATAACAGGCGTTTCCCTTGGAAAGCTCCCTCGTGCGCTCTCTGTTC 3792
QY 4194 CGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGGTGGCGCTTT 4253
Db 3793 CGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTGGGAAGGTGGCGCTTT 3852
QY 4254 CTCATGTCTCAGCTGTAGGTATCTCAGTTCCGTGTAGTGTGCTCGCTCAAGCTGGGCT 4313
Db 3853 CTCATGTCTCAGCTGTAGGTATCTCAGTTCCGTGTAGTGTGCTCGCTCAAGCTGGGCT 3912
QY 4314 GTGTGCAAAACCCCGCTTACCGCGGACCGCTGCGCTTATCCGGTAATACTATCGCTTG 4373
Db 3913 GTGTGCAAAACCCCGCTTACCGCGGACCGCTGCGCTTATCCGGTAATACTATCGCTTG 3972
QY 4374 AGTCCAAACCGGTAGACAGACTTATCGCACTGGCAGCAGCCACTGGTACAGATTAA 4433
Db 3973 AGTCCAAACCGGTAGACAGACTTATCGCACTGGCAGCAGCCACTGGTACAGATTAA 4032
QY 4434 GCAGAGCAGGTATGTAGGCGGTGTACAGAGTCTTTGAAGTGGTGGCCCTAACTACGGCT 4493
Db 4033 GCAGAGCAGGTATGTAGGCGGTGTACAGAGTCTTTGAAGTGGTGGCCCTAACTACGGCT 4092
QY 4494 AACTAGAAAGACAGTATTTGGTATCTGCTGTAAGCCAGTTACTTCGGAAAAA 4553
Db 4093 AACTAGAAAGACAGTATTTGGTATCTGCTGTAAGCCAGTTACTTCGGAAAAA 4152
QY 4554 GAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTGTGTTT 4613
Db 4153 GAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTGTGTTT 4212
QY 4614 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 4673
Db 4213 GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTCTA 4272
QY 4674 CCGGCTGAGCCGCTCAGTGGAAACGAAACTCACTTAAAGGATTTTGGTCATGAGATTAT 4733
Db 4273 CCGGCTGAGCCGCTCAGTGGAAACGAAACTCACTTAAAGGATTTTGGTCATGAGATTAT 4332
QY 4734 CAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 4793
Db 4333 CAAAAGGATCTTCACTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 4392
QY 4794 GTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAGGACCTATCT 4853
Db 4393 GTATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAGGACCTATCT 4452
QY 4854 CAGCATCTGTCTATTTTCTGTTTCAATAGTTGCTGATCCCGCTGCTGTGTAGTAACTA 4913
Db 4453 CAGCATCTGTCTATTTTCTGTTTCAATAGTTGCTGATCCCGCTGCTGTGTAGTAACTA 4512

```
Qy 4914 CGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCTCAATGATATACCGCGAGACCCAGCT 4973
Db 4513 CGATACGGGAGGCTTACCATCTGGCCCCAGTGTGCTCAATGATATACCGCGAGACCCAGCT 4572
Qy 4974 CACGGCTCCAGATTTATCAGCAATAAACACGAGCCGCGGAAGGCGGAGCGAGAGTG 5033
Db 4573 CACGGCTCCAGATTTATCAGCAATAAACACGAGCCGCGGAAGGCGGAGCGAGAGTG 4632
Qy 5034 GTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAA 5093
Db 4633 GTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAA 4692
Qy 5094 GTAGTTCCGAGTAATAGTTTGGCGAACGTTGTGCCATTTGCTACAGGCATCGTGTGT 5153
Db 4693 GTAGTTCCGAGTAATAGTTTGGCGAACGTTGTGCCATTTGCTACAGGCATCGTGTGT 4752
Qy 5154 CAGCTCGTGTGTGTATGGCTTCATTGAGCTCCGCTCCCAACGATCAAGCGAGTTA 5213
Db 4753 CAGCTCGTGTGTGTATGGCTTCATTGAGCTCCGCTCCCAACGATCAAGCGAGTTA 4812
Qy 5214 CATGATCCCGCATGTTGTGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTCA 5273
Db 4813 CATGATCCCGCATGTTGTGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTCA 4872
Qy 5274 GAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 5333
Db 4873 GAAGTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTTA 4932
Qy 5334 CTGTCTATGCCATCGGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCT 5393
Db 4933 CTGTCTATGCCATCGGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCT 4992
Qy 5394 GAGATAAGTGTATGCGCGCAGGTTGCTCTTCCCGCGGTCAATAACGGGATAATACCG 5453
Db 4993 GAGATAAGTGTATGCGCGCAGGTTGCTCTTCCCGCGGTCAATAACGGGATAATACCG 5052
Qy 5454 CGCCACATAGCAGAACTTTAAAAGTGTCTATCATTTGGAATAAGCTTTTCGGGGCGAAAC 5513
Db 5053 CGCCACATAGCAGAACTTTAAAAGTGTCTATCATTTGGAATAAGCTTTTCGGGGCGAAAC 5112
Qy 5514 TCTCAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGTGTGACCCCACT 5573
Db 5113 TCTCAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGTGTGACCCCACT 5172
Qy 5574 GATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAA 5633
Db 5173 GATCTTCAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAA 5232
Qy 5634 ATCCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTTT 5693
Db 5233 ATCCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTTT 5292
Qy 5694 TTCAATATTATGAGCAATTTATCAGGGTTATTGTCTCATGAGGGGATACATATTGAAAT 5753
Db 5293 TTCAATATTATGAGCAATTTATCAGGGTTATTGTCTCATGAGGGGATACATATTGAAAT 5352
Qy 5754 GTATTAGAAAAATAACAAATAGGGGTTCGGCGCACATTTCCCGAAAAAGTGCCACTG 5813
Db 5353 GTATTAGAAAAATAACAAATAGGGGTTCGGCGCACATTTCCCGAAAAAGTGCCACTG 5412
Qy 5814 ACGTCTAAGAAACCATTTATTATCATGACATTAACCTATATAAAATFAGCGGTATCACGAGC 5873
Db 5413 ACGTCTAAGAAACCATTTATTATCATGACATTAACCTATATAAAATFAGCGGTATCACGAGC 5932
Qy 5874 CTTTTCGTC 5882
Db 5473 CTTTTCGTC 5481
```

RESULT 10

US-09-132-808-1

; Sequence 1, Application US/09132808

; Patent No. 6197332

```
; GENERAL INFORMATION:
; APPLICANT: Ronald Zuckermann et al.
; TITLE OF INVENTION: Lipid-Conjugated Polyamide Compounds and Related
; TITLE OF INVENTION: Compositions and Methods Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,808
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1387.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4328 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-132-808-1
```

Query Match 36.8%; Score 2166.4; DB 3; Length 4328;
Best Local Similarity 70.4%; Pred. No. 0;
Matches 3060; Conservative 0; Mismatches 1236; Indels 52; Gaps 10;

```
Qy 333 GGAATATTGGCTATTGGCCATTGATAGCTTTGATCTATATCATATAATATGATATTA 392
Db 4 GCGGAATTTCCGATCTAGGCCATTGCATACGTTGTATCTATATCATATAATATGATATTA 63
Qy 393 TATTGCTCATGTCCTCATATGACCGCCATGTTGACATTTGATTTGACTAGTTATTAATA 452
Db 64 TATTGCTCATGTTCCCAATATGACCGCCATGTTGACATTTGATTTGACTAGTTATTAATA 123
Qy 453 GTAATCAATTTACGGGGTTCATTAGTTTCATAGCCCATATATGAGTTCGCGTTACATAACT 512
Db 124 GTAATCAATTTACGGGGTTCATTAGTTTCATAGCCCATATATGAGTTCGCGTTACATAACT 183
Qy 513 TAGCGTAATGGCCCGCTGGCTGACCGGCCAAGACACCCCGCCCATTTGACGTAAT 572
Db 184 TAGCGTAATGGCCCGCTGGCTGACCGGCCAAGACACCCCGCCCATTTGACGTAAT 243
Qy 573 GAGTATGTTCCCATAGTAACGCCAATAGGACTTTCCATTGACGTAATGGGTGGAGTA 632
Db 244 GAGTATGTTCCCATAGTAACGCCAATAGGACTTTCCATTGACGTAATGGGTGGAGTA 303
Qy 633 TTTTACGGTAATCGCCCACTTTGGCAGTACATCAAGTGTATCATATGCCAAGTCGCGCCCC 692
Db 304 TTTTACGGTAATCGCCCACTTTGGCAGTACATCAAGTGTATCATATGCCAAGTCGCGCCCC 363
Qy 693 TATTGACGTAATGACGGTAATGGCCGCTGGCATTTATGCCAGTATATGACGTAATGACG 752
Db 364 TATTGACGTAATGACGGTAATGGCCGCTGGCATTTATGCCAGTATATGACGTAATGACG 423
Qy 753 GGACTTTCTACTTGGCAGTACATCTACGTAATTTAGTCATCGCTATTACCATGTTGATGCG 812
Db 424 GGACTTTCTACTTGGCAGTACATCTACGTAATTTAGTCATCGCTATTACCATGTTGATGCG 483
```


Db 2605 ACTTGGCCCAATAGCAGCCAGTCCCTTCCGCTTTCAGTGACAAACGTCGAGCACAGCTGGC 2664
Qy 3033 GGGGATCTCCGCTGTCTTGGCAAAATTCGAGTCCCAACCTCCCAATCACTCACCA 3092
Db 2665 CAAGGAACGCGCTGCTGGCCAGCAGATAGCGGCGCTCCCTGCTTCTGCGAGTTCA --- 2721
Qy 3093 ACCTCCTGTCTCTCAATTTGCTCTGGTTATCGCTGGATGCTGCTGGCGGCTTTTATCAT 3152
Db 2722 -TTCAGGGCACGGACAGTCTGCTTTGACAAAAGAACCGGGCGCCCTGCGCTGACA 2779
Qy 3153 TTCTCTTCACTCTGCTGCTATGCTCATCTCTTATTTGTTCTTCTGATATCAAGGT 3212
Db 2780 GCCGGAACACGGCGCATCAGACAGCGGATGCTGTTGCTGCCAGTCAATACCGAATA 2839
Qy 3213 ATGTTGCGCTTGTCTCTTAATTCAGGATCAACAAACACAGTACGGGACCATGCAAA 3272
Db 2840 GCCTCTCCACCAGCGCGGAGAACCTCGTGCGAATCCATCTGTTCAATCATCGGAA 2899
Qy 3273 ACCTGCAAGACTCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCT 3332
Db 2900 ACATCTCTCATCTCTCTCTTG--ATCAGATCTTGATCCCTGGCCATCAGATCCTTG 2956
Qy 3333 ACCGATGGAATGCACTGTATCCCATCCATCGTCTGGGCTTTCCGCAAAATACCTA 3392
Db 2957 GCGGCAAGAAAGCCATCCAGTCTTACTTTGACGGGCTTCCCAACCTTACCAGAGGGGCC 3016
Qy 3393 TGGGAGTGGGCTCAGTCCGCTTCTCTTGGCTCAGTTTACTAGTGCAATTTGTTCAGTG 3452
Db 3017 CAGCTGGCAATTCGGTTCGCTTGCTGTCTCATAAACCGCCCATCTAGCTATCGGCATG 3076
Qy 3453 TTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAGCTATATGATGATGTTGGGG 3512
Db 3077 TAAGCCACTGCAAGCTACCTGCTTCTCTTGGCTTGGCTTTCCCTGTCAGATAG 3136
Qy 3513 CCAAGT--CTGTACAGATCGTGAATCCCTTTATACCGCTGTGTACCAATTTCTTTTGTG 3570
Db 3137 CCCAGTAGCTGACATTCACCGGGGTCAGCACCGCTTTCTCGGACTGGCTTTTACGTTG 3196
Qy 3571 TCTGGGTATACATTTAAGAAATCAGACTCGAGCAAGTCTAGAAAGCGCGCCAAATATC 3630
Db 3197 TCGCTTCTTTAGAGCCCTTCGCGCTCGAGTGTGCGGAGCGTGAAGCTGTCAAT 3256
Qy 3631 AAGGATCCACTACGCTGTAGAGCTCGCTGATCAGCCCTCGACTGCTCTTC--TAGTTGCC 3688
Db 3257 CCGGTTAAATTTTGTAAATCAGCTCATTTTAAACCAATAGGCGGAATCGCAAA 3316
Qy 3689 AGCATCTGTTGTTGGCCCTCCCGCTGCTTCTTTCAGCCCTGGAGGTCACCTCCCA 3748
Db 3317 TCCCTTATAAATCAAAAGAAATAGCCCGAGATAGGTTGAGTGTGTTTCCAGTTTGAACA 3376
Qy 3749 CTGTCCTTTCTTAATAAATGAGAAATTCATCGCATTTCTGAGTAGGTGTCATTCTA 3808
Db 3377 AGAGTCACTAATAAAGAGTGGACTCTCAACGTCAAAGGGCGGAAACCGCTATCAGG 3436
Qy 3809 TTCTGGGGGTGGGGTGGGCGAGCAGCAAGGGGAGGATTCGGGAAGACAATAGCAGGC 3868
Db 3437 GCATGCGGATCAGCTTATGCGGTGTGAATACCGACAGATCGCTAGGAGAAATAC 3496
Qy 3869 ATGCTGGGAGCTCTTCGCTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 3928
Db 3497 CGCATCAGGCGCTCTTCGCTCTCTGCTCACTGCTGCTGCTGCTGCTGCTGCTG 3556
Qy 3929 CGCGAGCGGTATCAGCTCACTCAAGGGCGGTAATACGGTTATCCACAGATCAGGGAT 3988
Db 3557 CGCGAGCGGTATCAGCTCACTCAAGGGCGGTAATACGGTTATCCACAGATCAGGGAT 3616
Qy 3989 AACCGAGGAAGAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAACCGTAAAGAGCC 4048
Db 3617 AACCGAGGAAGAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAACCGTAAAGAGCC 3676
Qy 4049 GCGTGTGCGGTTTTTCCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCAGCG 4108
Db 3677 GCGTGTGCGGTTTTTCCATAGGCTCCGCGCCCTGACGAGCATCAAAAATCAGCG 3736

Qy 4109 TCAAGTCAGAGTGGGCAAAACCGACAGACTATAAAGATACCAAGGCTTTCCCTCTGGA 4168
Db 3737 TCAAGTCAGAGTGGGCAAAACCGACAGACTATAAAGATACCAAGGCTTTCCCTCTGGA 3796
Qy 4169 AGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTT 4228
Db 3797 AGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTT 3856
Qy 4229 CTCCTTCCGGAGCGTGGGCTTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTG 4288
Db 3857 CTCCTTCCGGAGCGTGGGCTTTCTCATAGCTCAGCTGAGTATCTCAGTTCGGTG 3916
Qy 4289 TAGTTCGTTCCAAAGCTGGGCTGTGTCAGAACCCCTTTCAGGCCGACCGCTGC 4348
Db 3917 TAGTTCGTTCCAAAGCTGGGCTGTGTCAGAACCCCTTTCAGGCCGACCGCTGC 3976
Qy 4349 GCTTATCCGGTAACTATCTGTTAGTCCAAACCCGGTAAGACACGACTTATCGCACTG 4408
Db 3977 GCTTATCCGGTAACTATCTGTTAGTCCAAACCCGGTAAGACACGACTTATCGCACTG 4036
Qy 4409 GCAGCAGCCACTGGTAAAGGATTAAGCAGGAGGTATGAGGCGGTGCTACAGATTC 4468
Db 4037 GCAGCAGCCACTGGTAAAGGATTAAGCAGGAGGTATGAGGCGGTGCTACAGATTC 4096
Qy 4469 TTGAAGTGGTGGCTTAACTACGCTACACTAGAAAGCAGAGTATTTGGTATCTCGCTCTG 4528
Db 4097 TTGAAGTGGTGGCTTAACTACGCTACACTAGAAAGCAGAGTATTTGGTATCTCGCTCTG 4156
Qy 4529 CTGAAGCCAGTTACCTTCGGAAGAGTGTGATCTTTGATTCGGCAAAACCAACACC 4588
Db 4157 CTGAAGCCAGTTACCTTCGGAAGAGTGTGATCTTTGATTCGGCAAAACCAACACC 4216
Qy 4589 GCTGTAGCGGTGTTTTTTTTCGTCAGCAGCAGATTAACGGCAGAAAGAGTCT 4648
Db 4217 GCTGTAGCGGTGTTTTTTTTCGTCAGCAGCAGATTAACGGCAGAAAGAGTCT 4275
Qy 4649 CAAGAAGATCTTTGATCTTTTCTACGG 4676
Db 4276 CAAGAAGATCTTTGATCTTTTCTACTG 4303

RESULT 11
US-08-910-647-2
; Sequence 2, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706

; TELEFAX: (510) 655-3542		
; INFORMATION FOR SEQ ID NO: 2:		
; SEQUENCE CHARACTERISTICS:		
; LENGTH: 4328 base pairs		
; TYPE: nucleic acid		
; STRANDEDNESS: single		
; TOPOLOGY: linear		
; MOLECULE TYPE: DNA (genomic)		
US-08-910-647-2		
Query Match 36.8%; Score 2166.4; DB 3; Length 4328;		
Best Local Similarity 70.4%; Pred. No. 0;		
Matches 3060; Conservative 0; Mismatches 1236; Indels 52; Gaps 10;		
QY	333	GGGAATTATGGCTATTTGGCCATGTGACATACGTTGTATCTATATCAATAATATGACATTTA 392
DB	4	CGGGAATTTGCACTCTAGGCCAATGTCATACGTTGTATCTATATCAATAATATGACATTTA 63
QY	393	TATTGGCTCATGTCCCAATATGACCGCCATGTTGACATTTGATTTTACATTTTACATTTAATA 452
DB	64	TATTGGCTCATGTCCCAATATGACCGCCATGTTGACATTTGATTTTACATTTTATTAATA 123
QY	453	GTAAATCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCCGGGTTACATAACT 512
DB	124	GTAAATCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCCGGGTTACATAACT 183
QY	513	TACGGTAAATGGCCGCTGGCTGACCGCCCAACGACCCCGCCCAATTGACGTCATAATAAT 572
DB	184	TACGGTAAATGGCCGCTGGCTGACCGCCCAACGACCCCGCCCAATTGACGTCATAATAAT 243
QY	573	GACGTATGTTCCCATAGTAACGCCAATAGGAGCTTTCCATTGACGTCAAATGGGTGAGTA 632
DB	244	GACGTATGTTCCCATAGTAACGCCAATAGGAGCTTTCCATTGACGTCAAATGGGTGAGTA 303
QY	633	TTTACCGGTAAACTGGCCACTTTGGCAGTACATCAAGTGATATATGCCAAAGTCCGCCCCC 692
DB	304	TTTACCGGTAAACTGGCCACTTTGGCAGTACATCAAGTGATATATGCCAAAGTCCGCCCCC 363
QY	693	TATTGACGTCAAATGACGGTAAATGGCCGCTGGCCATTTATGCCCAGTACATGACCTTACG 752
DB	364	TATTGACGTCAAATGACGGTAAATGGCCGCTGGCCATTTATGCCCAGTACATGACCTTACG 423
QY	753	GGACTTTCTTACTTTGGCAGTACATCTACGTTATGATGATCATCGCTATTACCATGTTGATGG 812
DB	424	GGACTTTCTTACTTTGGCAGTACATCTACGTTATGATGATCATCGCTATTACCATGTTGATGG 483
QY	813	GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAAGTCT 872
DB	484	GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAAGTCT 543
QY	873	CCACCCCATTTGACGTCMAATGGGAGTTTGTGTCGACCAAAATCAACGGGACTTTCCAAA 932
DB	544	CCACCCCATTTGACGTCMAATGGGAGTTTGTGTCGACCAAAATCAACGGGACTTTCCAAA 603
QY	933	ATGTCGTAATAACCCCGCCGTTGACGAAATGGGCGGTAGCGGTGTTACGGTGGGAGGT 992
DB	604	ATGTCGTAATAACCCCGCCGTTGACGAAATGGGCGGTAGCGGTGTTACGGTGGGAGGT 663
QY	993	CTATATAAGCAGAGCTCGTTTATGTAACCGTTCAGATCGCTGAGAGCGCCATCCACGCTG 1052
DB	664	CTATATAAGCAGAGCTCGTTTATGTAACCGTTCAGATCGCTGAGAGCGCCATCCACGCTG 723
QY	1053	TTTTGACCTCCATAGAAGACACCGGACCGATCCAGCTCCGCGCGCGGGAGCGGTGCAT 1112
DB	724	TTTTGACCTCCATAGAAGACACCGGACCGATCCAGCTCCGCGCGCGGGAGCGGTGCAT 783
QY	1113	TGGAACCGCGATTTCCCGTCCCAAGAGTACGTAAGTACCGCTCATAGACTCTATAGGCA 1172
DB	784	TGGAACCGCGATTTCCCGTCCCAAGAGTACGTAAGTACCGCTCATAGACTCTATAGGCA 843
QY	1173	CACCCCTTTGGCTCTTATGATGCTACTACTGTTTGGCTTGGGCGCTATACACCCCGC 1232
DB	844	CACCCCTTTGGCTCTTATGATGCTACTACTGTTTGGCTTGGGCGCTATACACCCCGC 903
QY	1233	TCCTTATGCTATAGGTGATGATAGCTTAGCCTATAGCTGTGGTTATTGACCATTTATT 1292
DB	904	TCCTTATGCTATAGGTGATGATAGCTTAGCCTATAGCTGTGGTTATTGACCATTTATT 963
QY	1293	GACCACTCCCTATTGGTGACGATACCTTCCCAATTAATAATCCATACATGCGCTCTTTGCC 1352
DB	964	GACCACTCCCTATTGGTGACGATACCTTCCCAATTAATAATCCATACATGCGCTCTTTGCC 1023
QY	1353	ACAACATATCTTATTGGCTATATGCCAATACTCTGTCTTCCAGAGACTGACACGGACTCT 1412
DB	1024	ACAACATATCTTATTGGCTATATGCCAATACTCTGTCTTCCAGAGACTGACACGGACTCT 1083
QY	1413	GTATTTTACAGATGGGGTCCATTATTATTACAAATTCACATATACAAACGCCGT 1472
DB	1084	GTATTTTACAGATGGGGTCCATTATTATTACAAATTCACATATACAAACGCCGT 1143
QY	1473	CCCCGTGCCCGCAGCTTTTATTAAACATAGCGTGGGATCTCCGACATCTCGGCTACGTG 1532
DB	1144	CCCCGTGCCCGCAGCTTTTATTAAACATAGCGTGGGATCTCCGACATCTCGGCTACGTG 1203
QY	1533	TTCCGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCCA 1592
DB	1204	TTCCGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGTGCCA 1263
QY	1593	TCCGTCCAGCGCTCATGTGCTCGGAGCTCTTGTCTTAAACAGTGGAGGCCAGACT 1652
DB	1264	TCCGTCCAGCGCTCATGTGCTCGGAGCTCTTGTCTTAAACAGTGGAGGCCAGACT 1323
QY	1653	TAGGCACACACAATGCCCAACACACAGTGTGCCGACAAAGGCCGTGGCGTAGGGTA 1712
DB	1324	TAGGCACACACAATGCCCAACACACAGTGTGCCGACAAAGGCCGTGGCGTAGGGTA 1383
QY	1713	TGTGTCTGAAATGAGCTCGGAGATTGGCTCGACCTCGACCTGGACGAGATGGAAGACTTAAG 1772
DB	1384	TGTGTCTGAAATGAGCTCGGAGATTGGCTCGACCTCGACCTGGACGAGATGGAAGACTTAAG 1443
QY	1773	GCAGCGGACAGAGATGACGAGCTGAGTGTGTGTTTCTGATAAGAGTACAGAGTA 1832
DB	1444	GCAGCGGACAGAGATGACGAGCTGAGTGTGTGTTTCTGATAAGAGTACAGAGTA 1503
QY	1833	ACTCCCGTTCCGCTGCTGTTAAACGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCTGTGCT 1892
DB	1504	ACTCCCGTTCCGCTGCTGTTAAACGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCTGTGCT 1563
QY	1893	GCCGCGCGCCACACAGATAATAGTCAAGACTAAACAGACTGTTCTTCCATGGGT 1952
DB	1564	GCCGCGCGCCACACAGATAATAGTCAAGACTAAACAGACTGTTCTTCCATGGGT 1623
QY	1953	CTTTTCTGAGTCAACGCTGCTCGAGAAATCAAGCAATCATGGATGCAATGAAGAGGG 2012
DB	1624	CTTTTCTGAGTCAACGCTGCTCGAGCTAA-----GAATTCAGACTCGA 1666
QY	2013	CTCTGT 2072
DB	1667	GCAAGTCTAGAAAGCCATGGATATCGATCCACTACCGCTTAGAGCTCGTGTAGCCT 1726
QY	2073	CACGTCAACCGGGGAGTCCCGGCCACATGTGTCTGTGATTTGTTAGCCCTCTTCGCACCA 2132
DB	1727	CGACTGTGCTTCTAGTTGCCAGCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1786
QY	2133	GGGCCCAACAGNACGTCCAGCTGATCAACACCGGAGTTGGCACCTCAATAGCAG 2192
DB	1787	CCCTGGAAGTGCACCTCCACTGTCTTTCTTAATAAATGAGGAATTTGCATTCGCAAT 1846
QY	2193	GCCCTGAACTGCAATGATAGCTCAACACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2252
DB	1847	GTCTG-----AGTAGGTGTCTTCTTCTTCTGTGGGGTGGGGTGGGGAGGACAGC 1896
QY	2253	AGATTCAACTTTTACGGCTGTCTGAGAGCTAGCCAGCTGCCGACCCCTTACCGATTTT 2312
DB	1897	AAGGGGGAGATTGGGAAGACAATAGCAGGGGGTGGGGGAAGAACTCCAGCATGAGATC 1956

QY 2313 GACCAGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACCAAGCGCCCTAC 2372
Db 1957 CCCGCGTGGAGGATCATCAGCGCGGTCCCGGAAAGAGTTCCGGAAGCCCAACCTTTC 2016
QY 2373 TGTGGGACATACCCCCCAAAACCTTGGCGTATTGTGCCCCGGAAGAGTGTGTGTGCG 2432
Db 2017 ATAGAAGCGCGGTGGAA-----TCGAAATCTCGTGATGCAGGTTGGGCGTGGT 2072
QY 2433 GTATATTGCTTCACTCCAGCCCGTGTGTGGGAAGCACCGACGAGTGGGCGCGCC 2492
Db 2073 CGGTCAATTCGAACCCAGAGTCCGCTCAGAAGAACTCGTCAAGAGCGGATAGAAGC 2132
QY 2493 ACCTACAGCTGGGGTGAATATGATACGAGCTCTTGTCTTAAACATACAGGCCACCG 2552
Db 2133 GATGCGTGCGAATCGGGAGCGGATACCGTAAAGCAAGAGGAGCGGTAGCCCATTC 2192
QY 2553 CTGGGCAATTTGGTTGGTTGACTGTGATGAATCACTAAGTATCACAAGTGTGGGA 2612
Db 2193 GCGGCCAAGCTCTTCAGCAATATCAGCGGTAGCCACGCTATGTCTGTATAGCGGTCCG 2252
QY 2613 GCGCTCTTGTGTATCGGAGGGGGGCAACACACCCCTGCACCTGCCCACTGATTGC 2672
Db 2253 CACACCCAGCGGCCACAGTCGATGAATCCAGAAAGCGGCCAATTTCCACCATGAT--A 2310
QY 2673 TTCGCAAGCATCCGACGCCACATACTCTCGGTGGGCTCCGGTCCCTGGATCACACC 2732
Db 2311 TTCGCAAGCAGCATCGCCATGGGTACGACGAGATCTTCGCGTGGGCAATCGCGCC 2370
QY 2733 AGTGCTGTGTGACTACCCGTATAGGCTTTGGCATTTATCCTTGTATCAATCAATACCC 2792
Db 2371 TTGAGCTGTGCAACAGTTCGGGTGGCGGAGCCCTGTATGTCTTGTCTCGATCATCC 2430
QY 2793 ATATTAAATCAGATGATGAGTGGAGGGTGGACACAGGCTGGAGTGGCTGCAAC 2852
Db 2431 TGATCGAAGAGCGGCTTCATCCGAGTAGCTCGCTCGATGCGATTTTCGCTGG 2490
QY 2853 TGGACGGGGCGGAACCTTTCGATCTGGAAGATAGGACAGGTCCGAGATCGATATGGAG 2912
Db 2491 TGGTCGAATGGCAGGTAGCGG-----ATCAAGCGTATGCAGCGCGCATTCATCA 2544
QY 2913 AATCATCATCAGATTTCCTAGAACCCCTGCTCGTGTGTAAGGGGGGTTTTCTTGTG 2972
Db 2545 GCCATGATGATATCTTCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTCTGCCCGGC 2604
QY 2973 ACAAGATCTCTACATACCGCAGAGTCTAGATCGTGTGGAGTCTCTCAATTTCTA 3032
Db 2605 ACTTGGCCCAATAGCAGCCAGTCCCTTCCGCTTCAGTGACACGTCGAGCACAGTGG 2664
QY 3033 GGGGGATCTCCCGTGTGTCTTGGCCAAATTCGCAGTCCCAACCTCCATCACTCACCA 3092
Db 2665 CAAGGAAGCCCGTGGGCCAGCCAGATAGCGCGCTGCTCGTCTCGCAGTTCA--- 2721
QY 3093 ACCTCTGTCTCAATTTGTCTGGTTATCGCTGGATGTGTGGGGGTTTTTATCAT 3152
Db 2722 --TTCAGGGCACCGGACAGTCTGTGACAAAGAAACCGGGCGCCCTGGCGCTGACA 2779
QY 3153 TTCTCTTCTCTGCTGCTATGCTCTATCTTATTGTTCTTCTGGATTTCAAGGT 3212
Db 2780 GCCGGAACACGGGGCATCAGAGCAGCGATGTGTGTGGCCAGTCATAGCCGAATA 2839
QY 3213 ATGTTGCGCTTTGTCTCTAAATTCAGAGTCAACAAACACAGTACGGGACCATGCAAA 3272
Db 2840 GCCTCTCACCAAGCGCGGAGAACCTGGTGCATCATCTTGTTCATCATGCGAA 2899
QY 3273 ACCTGACGACTCTGCTCAAGGCACTCTATGTTTCCCTCATGTTGCTGTACAAACCT 3332
Db 2900 ACGATCTCATCTGCTCTTGT---ATCAGATCTTGATCCCTGCGGCATCAGATCCTG 2956
QY 3333 ACGATGGAATTTGCACCTGTATCCCATCCCATCTGCTGGGCTTTCGCAAAATACCTA 3392
Db 2957 GCGGCAAGAACCCATCAGTTTACTTTGAGGGCTTCCCAACCTTACAGAGGGGCGCC 3016
QY 3393 TGGAGTGGGCGCTCAGTCCCGTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAAGTGG 3452

Db 3017 CAGCTGGCAATTCGGGTTGCTGTGCTGCATAAAACCGCCAGTCTAGCTATCGCATG 3076
QY 3453 TTGCTAGGCGCTTCCCGCACTGTTTGGCTTTCAGCTATATGATGATGTTGGTATGGGG 3512
Db 3077 TAAGCCCATGCAAGTACCTGCTTCTCTTTGCGGCTGCGTTTCCCTTGTCCAGATAG 3136
QY 3513 CCAAGT--CTGTACAGCATCGTGGTCCCTTTTATACCGTGTTCACCAATTTCTTTTGTG 3570
Db 3137 CCAGTAGCTGACATTCATCCGGGTGAGCACCGCTTCTCGGACTGGCTTCTACGTGT 3196
QY 3571 TCTGGGTATACATTTAAGAAATTCAGACTCGAGCAAGTCTAGAAGGGCGCCAGATATC 3630
Db 3197 TCGCTTCTCTTAGCAGCCCTTGGCGCTGAGTGTCTGGCGCAGCGTGAAGCTGTCAAT 3256
QY 3631 AAGGATCCACTAGCGGTAGAGCTCGTGCATCAGCTCGACTGTGCTTC--TAGTTGCC 3688
Db 3257 CCGGTTAAATTTTGTAAATCAGCTCATTTTAAACCAATAGGCGCAAAATCGGCAAA 3316
QY 3689 AGCATCTGTGTTTGGCCCTCCCGCTGCTTCTTACCCCTGGAAGGTGCCACTCCCA 3748
Db 3317 TCCCTTATAATCAAAAGAAATAGCCCGAGATAGGTTGAGTGTGTTCCAGTTTGAACA 3376
QY 3749 CTGCTCTTCTTAATAAATGAGAAATTCATPCGCAATGTCTGAGTAGGTGTCATTCTA 3808
Db 3377 AGAGTCCACTATTAAAGAACTGGAATCCAAACGTCAAAGGGCGAAACCGTCTATCAGG 3436
QY 3809 TTTCTGGGGGTGGGTGGGCGAGCAGCAAGGGGAGGATTTGGGAAGACAATAGCAGGC 3868
Db 3437 GCGATGGCGGATCAGCTTATGCGGTGTGAATAACCGCACAGATGCGTAGAGGAAATAC 3496
QY 3869 ATGCTGGGAGCTTCTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTG 3928
Db 3497 CGCATCAGGGCTTCTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTG 3556
QY 3929 CGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGAT 3988
Db 3557 CGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGAT 3616
QY 3989 AAGCAGGAAGAACTGTGACAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGGCC 4048
Db 3617 AAGCAGGAAGAACTGTGACAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGGCC 3676
QY 4049 GCGTCTCGCGTCTTCTCATAGCTCCGCCCCCTGACGAGCATCACAAAATTCGACGC 4108
Db 3677 GCGTCTCGCGTCTTCTCATAGCTCCGCCCCCTGACGAGCATCACAAAATTCGACGC 3736
QY 4109 TCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCTTTTCCCTCGGA 4168
Db 3737 TCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATAACAGGCGTTTCCCTCGGA 3796
QY 4169 AGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTT 4228
Db 3797 AGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTT 3856
QY 4229 CTCCCTTCGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGTTG 4288
Db 3857 CTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTG 3916
QY 4289 TAGTCTGTTCCAGTCTGCGCTGTGTGACGAAACCCCGCTTACGCCCGACCGCTGC 4348
Db 3917 TAGTCTGTTCCAGTCTGCGCTGTGTGACGAAACCCCGCTTACGCCCGACCGCTGC 3976
QY 4349 GCCTTATCCGTAACCTATGCTTGTAGTCCAAACCGGTAGACACGACTTATGCCACTG 4408
Db 3977 GCCTTATCCGTAACCTATGCTTGTAGTCCAAACCGGTAAAGACACGACTTATGCCACTG 4036
QY 4409 GCAGCAGCCACTGCTGAACAGGATTTAGCAGAGCAGGTATGTAGCGGTGTCTACAGAGTTC 4468
Db 4037 GCAGCAGCCACTGCTGAACAGGATTTAGCAGAGCAGGTATGTAGCGGTGTCTACAGAGTTC 4096
QY 4469 TTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGGCTCTG 4528

Db 1204 TTCGGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCA 1263
Qy TCCGTCCAGCGGCTCATGGTTCGCTCGGAGCTCCTTCTCTTAACAGTGGAGCCAGACT 1652
Db TCCGTCCAGCGGCTCATGGTTCGCTCGGAGCTCCTTCTCTTAACAGTGGAGCCAGACT 1323
Qy TAGGCACAGCAATATGCCACCAACCACTAGTGTGCCGCAACAAGCCGCTGGCGGTAGGGTA 1712
Db TAGGCACAGCAATATGCCACCAACCACTAGTGTGCCGCAACAAGCCGCTGGCGGTAGGGTA 1383
Qy TGTGTCTGAAATAGCTCCGAGATTTGGGCTCGCACTCGACAGCAGATGGAAGACTTAAG 1772
Db TGTGTCTGAAATAGCTCCGAGATTTGGGCTCGCACTCGACAGCAGATGGAAGACTTAAG 1443
Qy GCAGCGGACAGAAAGATGAGGAGCTGAGTGTGTTGTTATTTCTGATAAGAGTCAAGAGTA 1832
Db GCAGCGGACAGAAAGATGAGGAGCTGAGTGTGTTGTTATTTCTGATAAGAGTCAAGAGTA 1503
Qy ACTCCCGTTCCGGTGTCTTAAACGGTGGAGGAGTGTAGTCTGAGCAGTACTCGTTGCT 1892
Db ACTCCCGTTCCGGTGTCTTAAACGGTGGAGGAGTGTAGTCTGAGCAGTACTCGTTGCT 1563
Qy GCCCGCGGCCACAGACATATAGCTGACAGTAACTAAGCAATCATGATGCAATGAAGAGAGGG 1952
Db GCCCGCGGCCACAGACATATAGCTGACAGTAACTAAGCAATCATGATGCAATGAAGAGAGGG 1623
Qy CTTTCTGACAGTCAACGCTCGTGCACAGTAACTCAAGCAATCATGATGCAATGAAGAGAGGG 2012
Db CTTTCTGACAGTCAACGCTCGTGCACAGTAACTCAAGCAATCATGATGCAATGAAGAGAGGG 1666
Qy CTCTGCTGTGTGCTGTGTGTGGAGCAGTCTTCTGTTTCGCCCGCAGCTAGCGAAACC 2072
Db GCAAGTCTAGAAAGCCATGATATCGGATCCACTACGCGTTAGAGCTCGCTGATCAGCCT 1726
Qy CAGTCAACCGGGGAGTCCCGGCCACATGCTGTCTGGATTTGTAGCTCTCTCGCACCA 2132
Db CGACTGTGCTTCTAGTGTGCGAGCCATCTGTGTGTTGTCCTCCCGTGCTCTCTTGA 1786
Qy GCGCCAAAGCAGAACGTCAGCTGATCAACACCAAGCGGAGTTGGCACCTCAATAGCACG 2192
Db CCTGGAGGTGCATCCCACTGTCTTCTTAATAAATGAGAAATGGAATCGAAT 1846
Qy GCCTGAACTGCAATGATAGACCTCAACACCGGCTGTGTGGCAGGCTTTTCTATACCAC 2252
Db GTCTG-----AGTAGTGTCTTCTTCTGGGGGTGGGTGGGCGAGGACAGC 1896
Qy AAGTTCACTCTTCAGGCTCTCTGAGAGCTAGCAGCTGCGGACCCCTTACCGATTT 2312
Db AAGGGGAGGATTTGGGAAGACAATAGCAGGGGGGTGGGCGAAGAACTCCAGCATGATC 1956
Qy GACCAGGCTGGGCGCCTATCAGTTATGCCAAGCGGCGCCCGCACAGCGCCCTAC 2372
Db CCGCGCTGAGGATCATCAGCGGGGTCCGGAAGACGATTCGGAAGCCCAACCTTTC 2016
Qy TGTGTCACCTACCCCGCAACCTTGGGTATTTGTGCCCGCAAGAGTGTGTGTGTCG 2432
Db ATAGAAGCGCGGTGGAA---TCGAATCTCGTGTATGTCAGTGTGGGCGTCTGTGT 2072
Qy GTATATTTCTTACTCCAGCCCGGTGTGTGGGAAGACGACAGCTGCGGCGCGCC 2492
Db CGGTCAATTTCCAAACCCAGAGTCCCGCTCAGAAAGAACTCGTCAAGAAAGCGGATAGAAGGC 2132
Qy ACCTACAGCTGGGTGAAATGATACGAGCTCTTCTGCTTAACTAACATACAGCGCACCG 2552
Db GATGCGGTGCAATTCGGGACGGGGATACCGTAAAGCACAGGAAGCGGTACGCCATTC 2192
Qy CTGGGCAATTTGGTTCGGTGTACTCGATGAATCACTCGATTCACCAAGTGTGCGGA 2612
Db GCCGCCAAGCTCTTCAGACATATACGGGTAGCCACGCTATGTCCTGTAGCGGTCCGC 2252
Qy GCGCTCTTGTGTATCGAGGGGGGGGCAACAAACCTGTCACCTGCCCCCATGTATGC 2672
Db CACACCCAGCGGCCACAGTCGATGAATCCAGAAAAAGCGGCCATTTTCCACCATGAT--A 2310

Qy TTCCGCAAGCATCCGAGCCACATACCTCTCGGTGGGCTCCGGTCCCTGGATCACACCC 2673
Db TTCCGCAAGCATCCGAGCCACATACCTCTCGGTGGGCTCCGGTCCCTGGATCACACCC 2370
Qy AGTGTCTCGTTCGACTACCCGATATAGGCTTTGGCATTAATCTTTGTACCAATCAACAC 2792
Db TTAGCGTTCGCGAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCC 2430
Qy ATATTTAAATCAGGATGACGTTGGGAGGGTGCACACAGGCTGGAAGCTGCTCAAC 2852
Db TGATCGACAAGACCGGCTTCCATCCGAGTACGCTCGCTCGATCGATGTTTCGCTGG 2490
Qy TGACCGCGGGGGAAGCTTTCGATCTGGAAGATAGGAGACAGGTCCTCGAGATATGGAG 2912
Db TGGTCCAATGGCAGGTAGCCGG-----ATCAAGCTATGCAGCGCCGATTCGATCA 2544
Qy AACATCACATCAGGATTCCTAGACCCCTGCTGCTGTGTACAGCGGGGTTTTCGTTG 2972
Db GCCATGATGATACCTTCTCGGAGGACAAAGTGTAGATGACAGGAGATCCTGCCCGGC 2604
Qy ACAAGATCCTCACATACGCGAGAGTCTAGACTCGTGTGGGACTTCTCTCAATTTCTA 3032
Db ACTTGCCCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAAACGTCGAGCAGCTGCG 2664
Qy GGGGATCTCCGCTGTGTCTTGGCCAAAATTCGAGTCCCCAACTCCCAATCACTCACCA 3092
Db CAAGGAACCGCGCTCGTGGCCAGCACGATAGCCGCGCTGCTCTGTCAGTTCA--- 2721
Qy ACTCTCTGCTCTCAATTTGTCTGTTATCGCTGGATGTGTCTGCGGGTTTTTATCAT 3152
Db TTAGGCGCACCGGACAGTCTGCTTGACAAAAAGAACCGGGCGCCCTGCGCTGACA 2779
Qy TTCTCTTCTCATCTGCTGTATGCTCTATCTTCTTATTTGGTCTTCTGGAATATCAAGT 3212
Db GCCGGAACACGGCGCATCAGAGCAGCGAATGTCTGTGTGCCCATAGTCCAGCCGAATA 2839
Qy ATGTTCGCGGTTGTCTCTTAATTCAGGATCAACAAACCAACAGTACGGGACCATCAAAA 3272
Db GCCTCTCCACCACGCGCGGAGAACCTGCGTGCAATCCAATCTTGTTCATCGGAA 2899
Qy ACTGCAACGACTCTGCTCAAGGCACTATATGTTTCCCTCATGTTGTGTGTACAAACCT 3332
Db ACGATCTCATCTGCTCTCTG---ATCAGATCTTGCATCCCTGCGCCATCAGATCCTTG 2956
Qy ACGGATGGAATTTGCACCTGTATTCCTATCCATCCCTGCGGCTTTCGCAAAATACCTA 3392
Db GCGGCAAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCCAACCTTACAGAGGGCGCCC 3016
Qy TGGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGTTTCAGTGG 3452
Db CAGCTGGCAATTCGCGTTCTGCTGTGCTCCATAAAACCGCCCATCTAGCTATCGCATG 3076
Qy TTGCTAGGCGCTTCCCGCATGCTGTTGGCTTTTTCAGCTATATGATGATGTTGTTGGGG 3512
Db TAAGCCACTGCAAGCTACTGCTTCTCTTTCGCGTTCGCTTTTCCCTTGTCCAGATAG 3136
Qy CCAAGT--CTGTACAGCATCGTGTGCTTATACCGCTGTGTACCAATTTCTTTTGTGTC 3570
Db CCAGTAGCTGACATTCATCCGGGTCAGCAACGTTTCTGCGGACTGGGCTTCTAGCTGT 3196
Qy TCTGGGTATACATTTAAGAAATTCAGACTCGCAAGTCTAGAAAGCGCGCCAGATATC 3630
Db TCCGCTTCTTTAGCAGCCCTTGGCGCTGAGTGTCTGGGAGCTGAAGCTGTCAATTT 3256
Qy AAGGATCCACTACGGGTTAGAGCTCGCTGATCAGCCCTCGACTGTGCTTC--TAGTTGCC 3688
Db CCGCGTAAATTTTGTAAATCAGCTCATTTTTTAAACCAATAGGCGGAAATCGGCAAAA 3316
Qy AGGCATCTGTTGTTTGGCCCTCCCGCTGCTTCTTGTAGCCCTGGAAGGTCGACCTCCCA 3748
Db TCCCTTATAATCAAAAGAAATAGCCCGAGATAGGGTTGAGTGTGTTTCCAGTTTGAACA 3376

```
QY 3749 CTGTCCTTTCTTAATAAATGAGAAATTCGCATCGCATTTGTCTGAGTAGTGTCATTCTA 3808
Db 3377 AGAGTCCACTATTAAAGAAACGTGACTCCAAACGTCAAGGGCGAAAAACCGTCTATCAGG 3436
QY 3809 TTCTGGGGGTGGGGTGGGSCAGACAGCAAGGAGGGAGGATTGGGAAGACAATAGCAGGC 3868
Db 3437 GCATGGCGATCAGCTTATGCGGTGTGAATACCGCACAGATGCGTAGGAGAAATAC 3496
QY 3869 ATGCTGGGAGCTCTTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGTCGGCTG 3928
Db 3497 CGCATCAGGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGTCGGCTG 3556
QY 3929 CGCGGAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGAT 3988
Db 3557 CGCGGAGCGGTATCAGTCACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGAT 3616
QY 3989 AACGAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGAGGC 4048
Db 3617 AACGAGGAAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGAGGC 3676
QY 4049 GCGTTGCTGCGCTTTTTCATAGGCTCGCCCCCTCGAGAGCATCACAAAAATCGACGC 4108
Db 3677 GCGTTGCTGCGCTTTTTCATAGGCTCGCCCCCTCGAGAGCATCACAAAAATCGACGC 3736
QY 4109 TCAAGTCAGAGGTGGGAAACCCGACAGACTATAAGATACAGCGCTTCCCTCTGA 4168
Db 3737 TCAAGTCAGAGGTGGGAAACCCGACAGACTATAAGATACAGCGCTTCCCTCTGA 3796
QY 4169 AGCTCCTCGTGGCTCTCTGCTTCGACCCCTCGCTTACCGGATACCTGCGGCTTT 4228
Db 3797 AGCTCCTCGTGGCTCTCTGCTTCGACCCCTCGCTTACCGGATACCTGCGGCTTT 3856
QY 4229 CTCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAAGCTGTAGGTATCTCAGTTCGGTG 4288
Db 3857 CTCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAAGCTGTAGGTATCTCAGTTCGGTG 3916
QY 4289 TAGTTCGTTCCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCGACCGCTGC 4348
Db 3917 TAGTTCGTTCCGCTCCAGCTGGGCTGTGTCAGCAACCCCGCTTACGCCGACCGCTGC 3976
QY 4349 GCCTTATCCGGTAACATATCTGTTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTG 4408
Db 3977 GCCTTATCCGGTAACATATCTGTTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTG 4036
QY 4409 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTC 4468
Db 4037 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGCGGTGCTACAGAGTTC 4096
QY 4469 TTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGATTTGGTATCTGCGCTCTG 4528
Db 4097 TTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGATTTGGTATCTGCGCTCTG 4156
QY 4529 CTGAAGCAGTTACCTTCGGAAAAAGAGTTGGTGGTCTTTGATCCGGCAAAACAAACACC 4588
Db 4157 CTGAAGCAGTTACCTTCGGAAAAAGAGTTGGTGGTCTTTGATCCGGCAAAACAAACACC 4216
QY 4589 GCTGGTAGCGGTGGTTTTTTTGTTCAGACAGCAGATTACGCGCAGAAAAAAGGATCT 4648
Db 4217 GCTGGTAGCGGG-1TTTTTTGTTTTCAGAGCAGAGATTACGCGCAGAAAAAAGGATCT 4275
QY 4649 CAGAGAGTCTTTGATCTTTTCTACGG 4676
Db 4276 CAGAGAGTCTTTGATCTTTTCTACTG 4303
```

RESULT 13
US-09-620-260-1
; Sequence 1, Application US/09620260
; Patent No. 6569450
; GENERAL INFORMATION:
; APPLICANT: Ronald Zuckermann et al.
; TITLE OF INVENTION: Lipid-Conjugated Polyamide Compounds and Related
; Compositions and Methods Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/620,260
FILING DATE: 09-Oct-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1387.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4328 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-260-1

Query Match 36.8%; Score 2166.4; DB 4; Length 4328;
Best Local Similarity 70.4%; Pred. No. 0;
Matches 3060; Conservative 0; Mismatches 1236; Indels 52; Gaps 10;
QY 333 GGGAAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATATAATGTACATTTA 392
Db 4 GCGGAATTTTCGACTCTAGGCCATTGCATACGTTGTATCTATATCATATAATGTACATTTA 63
QY 393 TATTGGCTCATCTCCAAATATGACCCGCAATGTTGACATTTGACATTTGACTAGTTAATAA 452
Db 64 TATTGGCTCATCTCCAAATATGACCCGCAATGTTGACATTTGACATTTGACTAGTTAATAA 123
QY 453 GTAATCAATTCAGGGTCAATTAGTTTCATAGCCCATATATGGAGTTCGCGGTTACATAACT 512
Db 124 GTAATCAATTCAGGGTCAATTAGTTTCATAGCCCATATATGGAGTTCGCGGTTACATAACT 183
QY 513 TACGGTAAATGCGCCGCTGGCTGACCGCCCAACGACCCCGCCCATTCGACGTCATTAAT 572
Db 184 TACGGTAAATGCGCCGCTGGCTGACCGCCCAACGACCCCGCCCATTCGACGTCATTAAT 243
QY 573 GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCATTTGGTGAGTA 632
Db 244 GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCATTTGGTGAGTA 303
QY 633 TTTTACGGTAAACTGCGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCC 692
Db 304 TTTTACGGTAAACTGCGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCC 363
QY 693 TATTGACGTCATAGCAGGTAATGCGCCGCTGGCAATTTAGCCAGTATGACGCTTACG 752
Db 364 TATTGACGTCATAGCAGGTAATGCGCCGCTGGCAATTTAGCCAGTATGACGCTTACG 423
QY 753 GGACTTTTCCTACTTGGCAGTACATCTACGTAATTAGTCATCGCTATTACCATGCTGATGG 812
Db 424 GGACTTTTCCTACTTGGCAGTACATCTACGTAATTAGTCATCGCTATTACCATGCTGATGG 483
QY 813 GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGTACTCAGCGGGGATTTTCCAAAGTCT 872
Db 484 GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGTACTCAGCGGGGATTTTCCAAAGTCT 543

QY 873 CCACCCCAATTGAGCTCAATGGAGTTTGTGTCACCAAAAATCAACGGGACTTTCAAA 932
DB 544 CCACCCCAATTGAGCTCAATGGAGTTTGTGTCACCAAAAATCAACGGGACTTTCAAA 603
QY 933 ATGTCTGTAATAACCCCGCCCGCTTGACGCAAAATGGGCGGTAGCGGTGTACGGTGGAGGT 992
DB 604 ATGTCTGTAATAACCCCGCCCGCTTGACGCAAAATGGGCGGTAGCGGTGTACGGTGGAGGT 663
QY 993 CTATATAAGCAGAGCTGTTTGTAGTGAACCGTCAAGTCGCTGAGAGCGCCATCCAGCGCTG 1052
DB 664 CTATATAAGCAGAGCTGTTTGTAGTGAACCGTCAAGTCGCTGAGAGCGCCATCCAGCGCTG 723
QY 1053 TTTTGACCTCCATAGAAAGACACCGGACCGGATCCAGGCTCCGGCGCGGGAAGCGGTGCAT 1112
DB 724 TTTTGACCTCCATAGAAAGACACCGGACCGGATCCAGGCTCCGGCGCGGGAACGGTGCAT 763
QY 1113 TGGAAACCGGATTTCCCGCTGCCAAGAGTGAAGTAACTACCGCTATAGACTCTATAGGCA 1172
DB 784 TGGAAACCGGATTTCCCGCTGCCAAGAGTGAAGTAACTACCGCTATAGACTCTATAGGCA 843
QY 1173 CACCCCTTTGGCTCTTATGATGATGCTATGCTTACTGTTTGGCTTGGGCTATACACCCCGC 1232
DB 844 CACCCCTTTGGCTCTTATGATGATGCTATGCTTACTGTTTGGCTTGGGCTATACACCCCGC 903
QY 1233 TCCTTATGCTATAGGTGATGCTATGCTTATAGCTTATAGGTGTTTATGACCATTTAT 1292
DB 904 TCCTTATGCTATAGGTGATGCTTATAGCTTATAGCTTATAGGTGTTTATGACCATTTAT 963
QY 1293 GACACTCCCTTATGCTGATGCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1352
DB 964 GACACTCCCTTATGCTGATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1023
QY 1353 ACAACTATCTCTATGCTGATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1412
DB 1024 ACAACTATCTCTATGCTGATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 1083
QY 1413 GTATTTTACAGGATGGGTCCATTTATTTATTTTACAAATTTACATATACAAACACCGCT 1472
DB 1084 GTATTTTACAGGATGGGTCCATTTATTTATTTTACAAATTTACATATACAAACACCGCT 1143
QY 1473 CCCCCTGCGCGAGTTTTTATTAACATAGGTGGGATCTCGACATCTCGGGTACGTG 1532
DB 1144 CCCCCTGCGCGAGTTTTTATTAACATAGGTGGGATCTCGACATCTCGGGTACGTG 1203
QY 1533 TTCGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCA 1592
DB 1204 TTCGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCA 1263
QY 1593 TCCGTCCAGCGGCTCATGCTCGGCTCGGAGCTCTTGTCTTAACTAGTGGAGGCCAGACT 1652
DB 1264 TCCGTCCAGCGGCTCATGCTCGGCTCGGAGCTCTTGTCTTAACTAGTGGAGGCCAGACT 1323
QY 1653 TAGGCAAGCAAAATGCCACCAACCAAGTGTGCGCAACAGGCGGTGGCGGTAGGGTA 1712
DB 1324 TAGGCAAGCAAAATGCCACCAACCAAGTGTGCGCAACAGGCGGTGGCGGTAGGGTA 1383
QY 1713 TGTGTCTGAANAATGAGCTCGGAGATTTGGGCTCGCACTTGACAGCCAGATGGAACACTTAAG 1772
DB 1384 TGTGTCTGAANAATGAGCTCGGAGATTTGGGCTCGCACTTGACAGCCAGATGGAACACTTAAG 1443
QY 1773 GCAGCGCAGAAAGATGAGGCTGAGTTGTTGTTATTTCTGATAAGAGTCAGAGGTA 1832
DB 1444 GCAGCGCAGAAAGATGAGGCTGAGTTGTTGTTATTTCTGATAAGAGTCAGAGGTA 1503
QY 1833 ACTCCGCTTGGGTGCTGTTTAAACGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCT 1892
DB 1504 ACTCCGCTTGGGTGCTGTTTAAACGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCT 1563
QY 1893 GCCGGCGGCGCCACGACATATAGTGAACACTTAAGCAGACTGTTCTTCCATGGGT 1952
DB 1564 GCCGGCGGCGCCACGACATATAGTGAACACTTAAGCAGACTGTTCTTCCATGGGT 1623

QY 1953 CTTTTCTGCAGTCAACCGTCTGCAGCAATTTCAAGCAATCATGATGCAATGAAGAGAGGG 2012
DB 1624 CTTTTCTGCAGTCAACCGTCTGCAGCAATTTCAAGCAATCATGATGCAATGAAGAGAGGG 1666
QY 2013 CTCTGCTGTGTGCTGTGCTGTGTGAGCAGTCTTGTGTTTCGCCAGCGCTAGGAAACC 2072
DB 1667 GCAAGTCTAGAAAGCCATCGATATCGGATCCACTACGCGTTAGAGCTCGTGTAGCAGCT 1726
QY 2073 CAGCTCACCGGGGAGTCCGGCCACACTGTCTGTGTTGTTGTTAGCTCTCTCGCACCA 2132
DB 1727 CGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGCCTCCCGGCTGCTTCTCTGA 1786
QY 2133 GGCGCAAGCAGAACGTCAGCTGATCAACCAACGAGCGAGTTGGCAGCTCTCAATAGCAG 2192
DB 1787 CCCTGGAAGGTGCACTCCCACTGTCTTCTTAATAAATAGGAAATTTGCATTCGCAATT 1846
QY 2193 GCCTGAACTGCAATGATAGCTCAACACGCGGTGTTGGCAGGCTTCTTCTATCACCAC 2252
DB 1847 GTCTG-----AGTAGGTGTCATTCTATTTCTGGGGGTGGGTGGGGCAGGACAGC 1896
QY 2253 AGCTTCAACTCTTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTT 2312
DB 1897 AAGGGGAGGATTTGGGAAGCAATAGCAGGGGGTGGGCGAAGAACTCCAGATGAGATC 1956
QY 2313 GACGAGGCTGGGGCTTATCAGTTATGCCAACGGAAGCGGGCCCCGACGAGCGCCCCCTAC 2372
DB 1957 CCGCGCTGGAGGATCATCCAGCGGCGTCCCGGNAACGATTTCCGAGGCCCAACCTTTC 2016
QY 2373 TGTGTGCACTACCCCCCAAAACCTTCCGGTATTTGTCCCGCGGAAGAGTGTGTGTGTCG 2432
DB 2017 ATAGAAAGGCGGGTGGAA-----TCGAAATCTCTGTATGGCAGTTTGGGCTGCTGTGGT 2072
QY 2433 GTATATTTGCTTCACTCCAGCCCGGTGTTGGGAACGACACGACAGGTGGGCGGCGCC 2492
DB 2073 CGGTCAATTTCCGAACCCAGAGTCCCCTCAGAAGAACTCGTCAAGAGGGGATAGAAGC 2132
QY 2493 ACCTACAGCTGGGGTGAATAATGATACGAGCTCTGCTCTTAAACAATACACAGGCCACCG 2552
DB 2133 GATGCGCTCGAATCGGAGCGCGGATACCGTAAAGCAGGAGAGCGGTACGCCCATTC 2192
QY 2553 CTGGGCAATTTGTTTGGTTGTTACTTGATGAATCACTGATGATTCACAAAGTGTGGGA 2612
DB 2193 GCGCCAAAGCTCTTACAGCAATATACGGGTAGCAACGCTATGCTCTGTAGAGCGGTCCG 2252
QY 2613 GCGCTCTTGTGTCATCGAGGGGCGGCAACCAACCTGACCTGCCCCACCTGATTCG 2672
DB 2253 CACACCCAGCGCGGCAAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGAT--A 2310
QY 2673 TTTCCGCAAGCATCCGAGCCACATATCTCTCGGTGGGCTCCGGTCCCTGGATCACACCC 2732
DB 2311 TTTCCGCAAGCATCCGAGCCATCGGTCACGACGAGATCTTCGCGGTGGGCGATCGCGCC 2370
QY 2733 AGTGTCTGTGCACTAACCGTATAGGCTTTGGCATTTATCTTTGTATACCACTCAACTACCC 2792
DB 2371 TTAGGCTCTGGCAAGTCTCGGTGGCGGAGCCCTGATGCTCTTCTGCTCAGATCATCC 2430
QY 2793 ATATTTAAATCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAAGCTGCTGCAAC 2852
DB 2431 TGATCACAAGACCGGCTTCCATCCGATGACGTCGCTCGGTGATGATGTTTTCGTTGG 2490
QY 2853 TGGACCGGGGGGAACGTTGCGATCTGGAAGATAGGGAAGGTCAGGATCCGAGATATGGAG 2912
DB 2491 TGGTCAATGGGCAAGTAGCCGG-----ATCAAGGATATGACGCGCGCATTTGCATCA 2544
QY 2913 AACATCAATCAGGATTTCTAGGACCCCTGCTGTTTACAGCGGGGTTTTTCTTGTGTTG 2972
DB 2545 GGCATGATGATACATTTCTCGGAGGCAAGGTGAGATACAGGAGATCTCTGCCCGGC 2604
QY 2973 ACAAGATCTCACAATACCGCAGAGTCTAGACTCGTGTGGGACTTCTCTCAATTTCTA 3032
DB 2605 ACTTCCGCAATAGCAGCAGTCCCTTCCGCTTTCAGTGACCAACGTCGAGCAGCTGG 2664
QY 3033 GGGGATCTCCGTTGTCTTGGCCAAAATTCGAGTCCCAACCTCCCACTCACTCACCA 3092

; LENGTH: 4328 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:			
US-09-620-259-1			
Query Match	36.8%;	Score 2166.4;	DB 4; Length 4328;
Best Local Similarity	70.4%;	Pred. No. 0;	
Matches 3060;	Conservative 0;	Mismatches 1236;	Indels 52; Gaps 10;
QY	333	GGGAATATTGGCTATTGGCCATTGCATAGCTGTGTATCTATATCATATAATATGACATTTA	392
DB	4	GCAGAAATTCGACTCTAGGCCATTGCATAGCTGTGTATCTATATCATATAATATGACATTTA	63
QY	393	TATTGGCTGATGTCATATGACGGCATGTTGACATTTGATTTAGCTAGTTATTAATA	452
DB	64	TATTGGCTCATGTCCAATATGACGGCATGTTGACATTTGATTTAGCTAGTTATTAATA	123
QY	453	GTAATCAATTTACGGGGTCAATTAGTTTCATAGCCCATATATGAGGTTCCGGGTTACATAACT	512
DB	124	GTAATCAATTTACGGGGTCAATTAGTTTCATAGCCCATATATGAGGTTCCGGGTTACATAACT	183
QY	513	TACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAAT	572
DB	184	TACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATAAT	243
QY	573	GACGTATGTTCCCATAGTAAGCCCAATAGGGAATTTCCATTTGACGTCAATGGGTGAGTA	632
DB	244	GACGTATGTTCCCATAGTAAGCCCAATAGGGAATTTCCATTTGACGTCAATGGGTGAGTA	303
QY	633	TTTACCGTAAACGTGCCACTTGGCAGTACATCAAGTGATCATATGCGAAGTCCGCCCC	692
DB	304	TTTACCGTAAACGTGCCACTTGGCAGTACATCAAGTGATCATATGCGAAGTCCGCCCC	363
QY	693	TATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTTATGCCAGTACATGACCTTACG	752
DB	364	TATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTTATGCCAGTACATGACCTTACG	423
QY	753	GGATTTCTTACTTGGCAGTACATCTACGTATTAGTATCGCTATTACCATGATGATGCG	812
DB	424	GGATTTCTTACTTGGCAGTACATCTACGTATTAGTATCGCTATTACCATGATGATGCG	483
QY	813	GTTTTGGCAGTACACCAATGGGGGTGATACGGGTTTGACTACGGGATTTCCAAAGTCT	872
DB	484	GTTTTGGCAGTACACCAATGGGGGTGATACGGGTTTGACTACGGGATTTCCAAAGTCT	543
QY	873	CCACCCCATTTGACGTCAATGGGAGTTTGTGTTGGCACCAAAATCAACGGGACTTTCCAAA	932
DB	544	CCACCCCATTTGACGTCAATGGGAGTTTGTGTTGGCACCAAAATCAACGGGACTTTCCAAA	603
QY	933	ATGTCGTATAAACCCCGCCCGTTGACGAAATGGGGCGTAGCGGTGTACGGTGGAGGT	992
DB	604	ATGTCGTATAAACCCCGCCCGTTGACGAAATGGGGCGTAGCGGTGTACGGTGGAGGT	663
QY	993	CTATATAAGCAGAGCTCGTTTGTAGTGAACCGTCAAGTCGCTGAGACGCCATCCACGCTG	1052
DB	664	CTATATAAGCAGAGCTCGTTTGTAGTGAACCGTCAAGTCGCTGAGACGCCATCCACGCTG	723
QY	1053	TTTGTGACCTCCATAGAGAACAACGGGACCGATCCAGCTCCGGCGGGGAAACGGTGCAT	1112
DB	724	TTTGTGACCTCCATAGAGAACAACGGGACCGATCCAGCTCCGGCGGGGAAACGGTGCAT	783
QY	1113	TGAAACCGGATTTCCCGTGCAGAGTGAAGTACCGCTATAGCTCTATAGGCA	1172
DB	784	TGAAACCGGATTTCCCGTGCAGAGTGAAGTACCGCTATAGCTCTATAGGCA	843
QY	1173	CACCCCTTTGGCTTTATGATGCTACTACTGTTTGGCTTGGGCTTATACACCCCGC	1232
DB	844	CACCCCTTTGGCTTTATGATGCTACTACTGTTTGGCTTGGGCTTATACACCCCGC	903
QY	1233	TCCTTATGCTATAGGTGATGGTATAGCTTAGCCTATAGGTGTGGGTTATTGACCATTTAT	1292
DB	904	TCCTTATGCTATAGGTGATGGTATAGCTTAGCCTATAGGTGTGGGTTATTGACCATTTAT	963
QY	1293	GACCACTCCCTATTGGTGACGATCTTTCCATTACTTAATCCATAACATGGCTCTTTGCC	1352
DB	964	GACCACTCCCTATTGGTGACGATCTTTCCATTACTTAATCCATAACATGGCTCTTTGCC	1023
QY	1353	ACAACTATCTCTATTGGCTATATGCAATACTCTGTCTCTTCCAGAGACTGACACGGACTCT	1412
DB	1024	ACAACTATCTCTATTGGCTATATGCAATACTCTGTCTCTTCCAGAGACTGACACGGACTCT	1083
QY	1413	GTAATTTTACAGATGGGGTCCATTATTATTTTCAAAATTCACATATACAAACGCCGT	1472
DB	1084	GTAATTTTACAGATGGGGTCCATTATTATTTTCAAAATTCACATATACAAACGCCGT	1143
QY	1473	CCCGCGTCCCGCAGTTTTTTTAAACATAGCGTGGGATCTCCGACATCTCGGGTACGTG	1532
DB	1144	CCCGCGTCCCGCAGTTTTTTTAAACATAGCGTGGGATCTCCGACATCTCGGGTACGTG	1203
QY	1533	TTCCGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCA	1592
DB	1204	TTCCGGACATGGGCTCTTCTCCGGTAGCGCGGAGCTTCCACATCCGAGCCCTGGTCCCA	1263
QY	1593	TCCGTCACGGCTCATGTGCTCGGACGTCTCTTGTCTTAAACATAGTGAGGCCAGACT	1652
DB	1264	TCCGTCACGGCTCATGTGCTCGGACGTCTCTTGTCTTAAACATAGTGAGGCCAGACT	1323
QY	1653	TAGGCACACGACAAATGCCCCACACACAGTGTGCCGACACAGGCCGTGGCGTAGGGTA	1712
DB	1324	TAGGCACACGACAAATGCCCCACACACAGTGTGCCGACACAGGCCGTGGCGTAGGGTA	1383
QY	1713	TGTGTCGTAATAATGAGCTCGGAGATTGGGCTCGACCTCGACATCGAGCAGATGGAAGACTTAAG	1772
DB	1384	TGTGTCGTAATAATGAGCTCGGAGATTGGGCTCGACCTCGACATCGAGCAGATGGAAGACTTAAG	1443
QY	1773	GCAGCGGACGAGAAAGATGCAGCAGCTGAGTTGTTGTTATTTCTGATAAGTGCAGAGGTA	1832
DB	1444	GCAGCGGACGAGAAAGATGCAGCAGCTGAGTTGTTGTTATTTCTGATAAGTGCAGAGGTA	1503
QY	1833	ACTCCCGTTGCGGTGCTGTTTAAACGTTGGAGGCGAGTGTAGTCTGACGAGTACTCTGTGCT	1892
DB	1504	ACTCCCGTTGCGGTGCTGTTTAAACGTTGGAGGCGAGTGTAGTCTGACGAGTACTCTGTGCT	1563
QY	1893	GCCGCGCGCGCCACACAGACATAATAGCTGACAGACTTAACAGACTGTTCTCTTCCATGGGT	1952
DB	1564	GCCGCGCGCGCCACACAGACATAATAGCTGACAGACTTAACAGACTGTTCTCTTCCATGGGT	1623
QY	1953	CTTTTCTGAGTCAACGCTGTCGACGAAATTAAGCAATCATGGATGCAATGAAGAGAGGG	2012
DB	1624	CTTTTCTGAGTCAACGCTGTCGACCTAA-----GAATTCAGACTCGA	1666
QY	2013	CTCTGCTGTGCTGCTGCTGTGTGGAGCAGTCTTCTGTTTGGCCCGAGCGCTAGCGAAACC	2072
DB	1667	GCAAGTCTAGAAAGCCATGATATCGGATCCATACGCTTAGAGCTCGCTGATCAGCCT	1726
QY	2073	CACGTCAACCGGGGAGTGCAGCCGACACTGTGTCTGAGATTTTGTAGCTCTCTCGCACCA	2132
DB	1727	CGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCGCGTCCCTTCTTGA	1786
QY	2133	GGCGCCAGAGAGAACGTCAGCTGATCAACCAACGGGAGTTGGGACCTCAATAGCAGC	2192
DB	1787	CCCTGGAAGTGCACCTCCCACTGCTCTTCTTAATAAATGAGGAAATTCGATCGCAT	1846
QY	2193	GCCTGAACTGCAATGATAGCCTCAACACCGGCTGTGCGAGGGCTTTTCTATCACCA	2252
DB	1847	GTCTG-----AGTAGGTGTCATTCTATTCTGGGGGTGGGGTGGGGCAGGACGC	1896
QY	2253	AAGTTCAACTCTTACGGCTGTCTGAGAGGCTAGCCAGCTGCGCAGCCCTTTACCGATTTT	2312
DB	1897	AAGGGGAGGATTTGGGAAGACAATAGCAGGGGGTGGCGAAGAACTCCAGCATGAGATC	1956
QY	2313	GACCAAGGCTGGGGCCCTATCAGTTTATGCCAACGGAAGGGGCCCGGACACGCGCCCTAC	2372

Db	1957	CCGCGCTGGAGGATCATCCAGCGCGCTCCCGGAAAAAGATTCGGAAGCCCAACCTTTC	2016
QY	2373	TGCTGGCACTACCCCCCAAAACCTTGGGTAATTGTGCCGCGGAAGTGTGTGTGTCGG	2432
Db	2017	ATAGAAGCGCGGTGGAA-----TCGAAATCTCGTATGGCAGGTTGGGCGTGCCTTGGT	2072
QY	2433	GTATATTGCTTCACTCCAGCCCGGTGGTGGGAACGACGACAGGTGCGGCGCGCC	2492
Db	2073	CGGTCAATTTGAAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGGGGATAGAAGC	2132
QY	2493	ACCTACAGCTGGGGTGAATAATGATACGGACGTCTTCGTCCTTAACAATACCAGGCCACCG	2552
Db	2133	GATGCGCTGCAATCGGAGCGCGGATACCGTTAAAGACAGAGAGCGGTACGCCCATTC	2192
QY	2553	CTGGGCAATTGTTGGTGTGTACTCGATGAATCAACTGGATTCACCAAGTGTGCGGA	2612
Db	2193	GCGCCAAAGCTCTTACGAATATACGGGTAGCAACGCTATGCTGTATAGCGGTCCGC	2252
QY	2613	GCCTCTCTGTGTCATCGAGGGCGGCAACAACACCTGTCACCTGCCCTGATTCG	2672
Db	2253	CACACCGCGGCCACAGTCGATGAATCCAGAAAGCGGCCAATTTTCCACCAATGAT--A	2310
QY	2673	TTCCGCAAGCATCCGGAACGACATATCTCTCGTGGGGTCCGGTCCCTGGGATCACACCC	2732
Db	2311	TTCCGCAAGCAGCATCGCATGGGTACGACGAGATCTCTGCCGTGGGCGATCGCGCC	2370
QY	2733	AGTGCTGTGTCACATACCGGTATAGGCTTTGGCATTTATCTTGTATACCATCAACTACCC	2792
Db	2371	TTGAGCTGTGCAACAGTTTCGGTGGCGGAGCCCTGTATGCTCTTCGTCCAGATCATCC	2430
QY	2793	ATATTAAATACAGGTATGATGAGTGGGAGGGGTGGAACACAGGCTGGAAGTGCCTCAAC	2852
Db	2431	TGATCGACAAGACCGGCTTCATCCGAGTACGTCTCGCTCGATGGATGTTTCGCTTGG	2490
QY	2853	TGACCGGGGCGAACTTTCGATCTGGAAGATAGGGAACAGGTCCGAGATTCGATATGGAG	2912
Db	2491	TGCTCGAATGGCGAGTAGCCGG-----ATCAAGCGTATGCAGCGCGCATTCGATCA	2544
QY	2913	AACATCACATCAGGATTCCTAGAACCCCTGCTCGTGTATCAGCGCGGGTTCCTTGTGG	2972
Db	2545	GCCATGATGATATCTTCTCGGAGGACGAAGGTGAGTGCACAGGAGATCTCGCCCGGC	2604
QY	2973	ACAAGAATCTCAATACCGCAGAGTCTAGACTCGTGTGGTGAATCTCTCAATTTCTA	3032
Db	2605	ACTTCGCCCAATAGACCGCATCCCTTCGCGTTTCTAGTGACAACGTCGAGCACAGTGGC	2664
QY	3033	GGGGGATCTCCCTGTGTCTTTGGCCAAAAATTCGACGTCCCAACCTCCCAATCACTCACCA	3092
Db	2665	CAAGGAAGCCCGTCTGTGGCCGCCACGATAGCGCGCTGCTCGTCTCTGCAAGTTCA---	2721
QY	3093	ACCTCTGTCTCAATTTGCTGTGTTATCGTGTGATGTCGCGGCGTTCATCATTA	3152
Db	2722	--TTCAGGGCACGGACAGTCCGCTCTTGAACAAAAAGAACCGGGCGCCCTGCGCTGACA	2779
QY	3153	TTCTCTCTTCACTGCTGCTATGCCATCTCTTCTTATGTTCTCTTGGATATCAAGGT	3212
Db	2780	GCCGGAACACGGCGCATCAGACGCGGATGCTGTGTGTCGCCAGTCAATAGCCGAATA	2839
QY	3213	ATGTTCGCCGTTTGTCTCTTAATTCAGGATCAACCAACCAACAGTACGGGACCATGCAAA	3272
Db	2840	GCCTCTCCACCAAGCGGCGGAGAACCTCGTGTGCAATCCATCTTGTTCATCATGCGAA	2899
QY	3273	ACCTGACAGACTCTGTCTCAAGGAACTCTATGTTTCCTTCATGTTGCTGTGTACAAACCT	3332
Db	2900	ACGATCTCATCTGTCTCTTG---ATCAGATCTTGATCCCTCGCGCCATCAGATCCCTTG	2956
QY	3333	ACGATGGAATTCGACCTGATTTCCCATCCCATCGTCTGGGCTTTTCGAAAAATACCTA	3392
Db	2957	GCGCAAGAAAGCCATCAGTTACTTTGAGGGCTTCCCAACCTTACACAGGGGCGCC	3016
QY	3393	TGGAGTGGGCCCTCAGTCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGG	3452
Db	3017	CAGCTGGCAATTCGGGTTCCGTTGCTGCTCCATAAACCGGCCAGTCTAGCTATCGGCATG	3076
QY	3453	TTCTGAGGCTTTCCCCCACTGTTTGGCTTTTCCAGCTATATGATGATGTTGGGCTT	3512
Db	3077	TAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGGCGTTCGCTTTTCCCTTGTCCAGATAG	3136
QY	3513	CCAAGT--CTGTACAGCATCGTGTAGTCCCTTTATATACCGCTGTACCAATTTTCTTTTGT	3570
Db	3137	CCAGTAGCTGACATTCATCCGGGTGACACCGTTTCTGCGGACTGGCTTTCTAGCTGT	3196
QY	3571	TTCTGGGTATACATTTAAGAAATTCAGACTCGAGCAAGTCTTAGAAAGCGCGCAAGATATC	3630
Db	3197	TCCGCTTCTTTAGCAGCCCTTTCGCGCTGTGTGTGCGCAGCGTGAAGCTGTCAAT	3256
QY	3631	AAGGATCCACTAGCGGTAGAGCTCGCTGATCAGCTCGACTGCTGCTTC--TAGTTGCC	3688
Db	3257	CCGCGTTAAATTTTGTTAATCAGCTCATTTTTTAAACCAATAGGCCGAAATCGGCAAA	3316
QY	3689	AGCCATCTGTTTGGCCCTCCCGCTTCTTGACCTTCCCTGGAAGGTGCACCTCCCA	3748
Db	3317	TCCCTTATAAATCAAAAGNATAGCCCGAGATAGGGTTGAGTGTGTTCCAGTTTGGAAACA	3376
QY	3749	CTGTCTTCTTCTTAATAAATGAGGAAATTCGATTCGATTTGTCTGAGTAGGTGTCTCTA	3808
Db	3377	ACAGTCCACTATTAAAGAACGTGGAATCCACGCTCAAAGGGCGAAAAACCGTCTATCAG	3436
QY	3809	TTCTGGGGGTGGGTGGGCGAGGACAGCAAGGGGAGGATTTGGGAAGACAATAGCAGC	3868
Db	3437	GCGATGGCGGATCAGCTTATGCGGTGTAAATACCGCACAGATGCGTAGGAGAAAATAC	3496
QY	3869	ATGCTGGGAGCTCTTCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTG	3928
Db	3497	CGCATCAGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGGCTG	3556
QY	3929	CGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGATCAGGGAT	3988
Db	3557	CGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGATCAGGGAT	3616
QY	3989	AACGAGGAAGAACATGTGACCAAGGCGCAGCAAAAGCGGAGAACCGTAAAAAGGCC	4048
Db	3617	AACGAGGAAGAACATGTGACCAAGGCGCAGCAAAAGCGGAGAACCGTAAAAAGGCC	3676
QY	4049	GGTGTCTGCGCTTTTTCATAGGCTCCGCCCTTCGACGAGCATCACAAAAATCGACGC	4108
Db	3677	GCTTGTCTGCGCTTTTTCATAGGCTCCGCCCTTCGACGAGCATCACAAAAATCGACGC	3736
QY	4109	TCAAGTCAGAGTGGCGAAACCCGACAGACTATATAAGATACAGGCGTTTCCCTCGGA	4168
Db	3737	TCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCTCGGA	3796
QY	4169	AGCTCCCTCGTGGCTTCTCTGTTCCGACCCCTTCGCGATACCGGATACCTGCTCCGCTTT	4228
Db	3797	AGCTCCCTCGTGGCTTCTCTGTTCCGACCCCTTCGCGATACCGGATACCTGCTCCGCTTT	3856
QY	4229	CTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACTGCTGAGTATCTCAGTTCGGTG	4288
Db	3857	CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACTGCTAGGTATCTCAGTTCGGTG	3916
QY	4289	TAGTGTGCTCGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCCCGACCGCTGC	4348
Db	3917	TAGTGTGCTCGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCCCGACCGCTGC	3976
QY	4349	GCCTTATCCGTTAACTATCTGTTTCCGACCCCTTAAGACACCACTTATCGCCACTG	4408
Db	3977	GCCTTATCCGTTAACTATCTGTTTCCGACCCCTTAAGACACCACTTATCGCCACTG	4036
QY	4409	GCAGCAGCACTCGTAAACAGGATTTAGCAGAGCGAGGTATGTAGGGCGTCTACAGAGTTC	4468
Db	4037	GCAGCAGCACTCGTAAACAGGATTTAGCAGAGCGAGGTATGTAGGGCGTCTACAGAGTTC	4096
QY	4469	TTGAAGTGTGGCTTAACCTACCGCTACACTAGAGSACAGTATTTGGTATCTGCGCTCTG	4528
Db	4097	TTGAAGTGTGGCTTAACCTACCGCTACACTAGAGSACAGTATTTGGTATCTGCGCTCTG	4156

Qy 4529 CTAAGCCAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACACACC 4588
Db CTAAGCCAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACACACC 4216
Qy 4589 GCTGGTAGCGGTGTTTGTGTCAGCAGCAGATTACGGCGCAGAAAAAGGATCT 4648
Db 4217 GCTGGTAGCGCGG- TTTTGTGTCAGCAGCAGATTACGGCGCAGAAAAAGGATCT 4275
Qy 4649 CAAGAAGATCCTTTGATCTTTCTACGG 4676
Db 4276 CAAGAGATCCTTTGATCTTTCTACTG 4303

RESULT 15
US-08-799-569-1
; Sequence 1, Application US/08799569
; Patent No. 6133244
; GENERAL INFORMATION:
; APPLICANT: Michel, Marie-Louise
; APPLICANT: Mancine, Maryline
; TITLE OF INVENTION: Nucleotide Vector, Composition
; TITLE OF INVENTION: Containing Such Vector, and Vaccine for Immunization
; TITLE OF INVENTION: Against Hepatitis
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,569
; FILING DATE: 12-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/706,337
; FILING DATE: 27-APR-1994
; APPLICATION NUMBER: 30-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/633,821
; FILING DATE: 22-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/00483
; FILING DATE: 27-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0128-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5618 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-799-569-1

Query Match 36.7%; Score 2156.6; DB 3; Length 5618;
Best Local Similarity 63.9%; Pred. No. 0;
Matches 3464; Conservative 0; Mismatches 1899; Indels 58; Gaps 11;
Qy 422 GTTGACATTGATTATGCTAGTATTATATAGTAATCAATTACGGGGTCAATTAGTTCAATA 481
Db 232 GTTGACATTGATTATGCTAGTATTATATAGTAATCAATTACGGGGTCAATTAGTTCAATA 291

Qy 482 GCCATATATGAGTTCCGCGTTTACATAACTTACGGTAAATGGCCGCGCTGGCTGACCGC 541
Db 292 GCCATATATGAGTTCCGCGTTTACATAACTTACGGTAAATGGCCGCGCTGGCTGACCGC 351
Qy 542 CCAACGACCCCGCCCATTTGACGTCAATTAATGACGTATGTTCCCATAGTAACGCCAATAG 601
Db 352 CCAACGACCCCGCCCATTTGACGTCAATTAATGACGTATGTTCCCATAGTAACGCCAATAG 411
Qy 602 GGACTTTTCATTGACGTCAATGGGTGGAGTATTACGGTAAATGGCCGCGCTGGCTGACCGC 661
Db 412 GGACTTTTCATTGACGTCAATGGGTGGAGTATTACGGTAAATGGCCGCGCTGGCTGACCGC 471
Qy 662 ATCAAGTGTATCATATGCCAAGTCCGCGCCCTTATTGACGTCAATGACGGTAAATGGCCCG 721
Db 472 ATCAAGTGTATCATATGCCAAGTCCGCGCCCTTATTGACGTCAATGACGGTAAATGGCCCG 531
Qy 722 CTTGGCATTTATGCCGAGTACATGACCTTACGGGACCTTCTTACTTGGCAGTACATCTACG 781
Db 532 CTTGGCATTTATGCCGAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG 591
Qy 782 TATTAGTCATCGCTATTACCATGGTGTATGGGTTTGGCAGTACACCAATGGCGGTGGAT 841
Db 592 TATTAGTCATCGCTATTACCATGGTGTATGGGTTTGGCAGTACATCAATGGCGGTGGAT 651
Qy 842 AGCGGTTTGACTCACGGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGT 901
Db 652 AGCGGTTTGACTCACGGGGATTTCCAAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGT 711
Qy 902 TTTGGCACAAATAACAAACGGGACTTTTCCAAAATGTCGTAATAACCCGCGCCCTTGGACGC 961
Db 712 TTTGGCACAAATAACAAACGGGACTTTTCCAAAATGTCGTAATAACCCGCGCCCTTGGACGC 771
Qy 962 AATGGCGGTAGGCGGTACGGTGGAGGTCTATATAGCAGAGCTCGTTAGTAGTAAAC 1021
Db 772 AATGGCGGTAGGCGGTACGGTGGAGGTCTATATAAGCAGAGCTCTTGGCTAACTA 831
Qy 1022 GTCAGATCGCTGGAGACGCCATCCACAGCTGTTTTCACCTCCATAGAAAGACACCGGAC 1081
Db 832 GAGAACCCACTGCTTAACTGGCTTATCGAAAT--AATACGACTCACTATAGGAGACC 888
Qy 1082 GATCCAGCTCCCGCGCGCGGAAACGGTGCATTGGAAACGGGAGTTCCTCCGTCGCAAGTG 1141
Db 889 CAAGCTTGGTACCGGGCGCCCTCTCAGGATTTGGG-----ACCTCGCTGAACA 939
Qy 1142 AGTAAGTACCGCTATAGACTCTATAGCAGACACCCCTTTGGCTCTTATGCTATATAC 1201
Db 940 TGGAGAACATCATCATCAGGATTCCTAGG-----ACCCCTTCTCGTGTACAGCGGGGTTT 995
Qy 1202 TGTTTTGGCTTGGGGCTTATACACCCCGCTCTTATGCTATAGGTGATGTTAGCTT 1261
Db 996 TTTCTTGTGACAAGAACTCTCAATACCGCAGAGTCTAGACTCGGTGGACTTCTCTC 1055
Qy 1262 AGCTATAGGTGGGTATTGACCAATATTGACCACTCCCTTATTTGGTGACGATCTTT 1321
Db 1056 AATTTTCTAGGGGAACTACCGGTGTCTTGGCAAAATTCGCAGTCCCAACCTCCAAT 1115
Qy 1322 CCATTACTAATCCATPAACATGGCTCTTTGCCACACTATCTCTATTTGGCTATATGCCAAT 1381
Db 1116 CACTACCAACCTCTGTCTCCTCAACTTGTCTCGGTATATCGTGGATGTCTGCGGCGT 1175
Qy 1382 ACTCTGTCTTCAGAGACTGACACGGACTCTGTAT-TTTTACAGGATGGGTCCATTAT 1440
Db 1176 TTTATCATCTTCTTCTCATCTCTGTGCTATGCTCATCTTCTTGTGGTCTTCTGGAC 1235
Qy 1441 TATTTACAAATTCACATATACAAACACCGCTCCCGCGCGAGTCTTTTATTAACA 1500
Db 1236 TATCAAGGTATGTTGCGCGCTTTGTCTCTAAATTCAGGATCTCTCAACCAACAGCAGGGA 1295
Qy 1501 TAGCGTGGGATCTCCGACATCTCGGTAGCTGTTTCCGACATGGGCTCTTCTCCGCTAGC 1560
Db 1296 CCATCGCGGACCTGCATGACTACTGCTCAAGAACTCTATGATATCCCTCTGTTGCTGT 1355

Db 3498 ACAATTCACACACATACAGAGCCGGAAGCATAAAGTGTAAAGCCTCGGGTGCCTAATGA 3557
Qy 3771 GGAAATTCGATCGCATTTCTAGTAGTGTGTCATTTCTATTCGGGGGCTGGGGTGG---G 3827
Db 3558 GTGAGCTAACTCACAATTAATTGGCTGGCTCACTGCCGCTTCCAGTTCGGNAACCTG 3617
Qy 3928 GCAGGACAGCAGAGGGGAGGATTTGGGAAGACAATAGCAGGCATGC-----TGGGG 3877
Db 3618 TCGTGCCAGCTGCATTAATGAATCGGCCAAACCGCGCGGAGAGCGGTTTGGCGTATTGGG 3677
Qy 3978 AGCTTTTCGCTTCCTCGCTCACTGACTCCGCTCGCTCGCTCGCTCGCTCGCTCGCGCAGCG 3937
Db 3678 CGCTTTTCGCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGCAGCG 3737
Qy 3938 GTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGGGATTAACGACGGA 3997
Db 3738 GTATCAGCTCACTCAAAAGCGGTAAATACGGTTATCCACAGATCAGGGGATTAACGACGGA 3797
Qy 3998 AAGAACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTTAAAGAGCCGGTTGCTG 4057
Db 3798 AAGAACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTTAAAGAGCCGGTTGCTG 3857
Qy 4058 GCCTTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAATCAGACGCTCAAGTCAG 4117
Db 3858 GCCTTTTTCATAGCTCCGCCCCCTGACGAGCATCAAAAATCAGACGCTCAAGTCAG 3917
Qy 4118 AGGTGGCGAAACCCGACAGACTATAAAGATACAGAGCGTTTCCCCCTGGAAGCTCCCTC 4177
Db 3918 AGGTGGCGAAACCCGACAGACTATAAAGATACAGAGCGTTTCCCCCTGGAAGCTCCCTC 3977
Qy 4178 GTGCGCTCTCCTGTTCCGACCCCTGCTTACCGGATACCGGATACCTGTCGCTTTTCCCTTCG 4237
Db 3978 GTGCGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTTTCCCTTCG 4037
Qy 4238 GGAAGCTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGCCTT 4297
Db 4038 GGAAGCTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGTGCCTT 4097
Qy 4298 CGCTCCAAAGCTGGCTGTGTGACGAACCCCGCTTACGAGCCGACCGCTCGCTTATCC 4357
Db 4098 CGCTCCAAAGCTGGCTGTGTGACGAACCCCGCTTACGAGCCGACCGCTCGCTTATCC 4157
Qy 4358 GGTAACTATCGTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCC 4417
Db 4158 GGTAACTATCGTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCC 4217
Qy 4418 ACTGTTAAAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG 4477
Db 4218 ACTGTTAAAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG 4277
Qy 4478 TGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA 4537
Db 4278 TGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA 4337
Qy 4538 GTTACCTTCGGAAAAAGAGTTGGTGTCTTGTATCGCGCAAAACAAACCCGCTGGTTCG 4597
Db 4338 GTTACCTTCGGAAAAAGAGTTGGTGTCTTGTATCGCGCAAAACAAACCCGCTGGTTCG 4397
Qy 4598 GGTGGTTTTTTTGTTCGACAGCAGCATTAACGGCAGAAAAAAGGATCTCAAGAAGAT 4657
Db 4398 GGTGGTTTTTTTGTTCGACAGCAGCATTAACGGCAGAAAAAAGGATCTCAAGAAGAT 4457
Qy 4658 CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAATCAGCTTAAAGGATT 4717
Db 4458 CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAATCAGCTTAAAGGATT 4517
Qy 4718 TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGT 4777
Db 4518 TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGT 4577
Qy 4778 TTTAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTTACCAATCGTTAATC 4837
Db 4578 TTTAAATCAATCTAAAGTATATAGTAAACTTTGGTCTGACAGTTTACCAATCGTTAATC 4637

Qy 4838 AGTGAGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCC 4897
Db 4638 AGTGAGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTCCCC 4697
Qy 4898 GTCGTGTAGATAACTACGATAACCGGAGGGCTTACCATCTCGGCCCCAGTGTGCAATGATA 4957
Db 4698 GTCGTGTAGATAACTACGATAACCGGAGGGCTTACCATCTCGGCCCCAGTGTGCAATGATA 4757
Qy 4958 CCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGG 5017
Db 4758 CCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGG 4817
Qy 5018 GCCGAGCGCAGAAAGTGGTCTCTGCAACTTTATCCGCGCTCCATCCAGTCTATTAATTTGTC 5077
Db 4818 GCCGAGCGCAGAAAGTGGTCTCTGCAACTTTATCCGCGCTCCATCCAGTCTATTAATTTGTC 4877
Qy 5078 CGGGAAGCTAGAGTAAGTAGTTCCGCAAGTAATAGTTTGGCAACCGTTGTCATTTGCT 5137
Db 4878 CGGGAAGCTAGAGTAAGTAGTTCCGCAAGTAATAGTTTGGCAACCGTTGTCATTTGCT 4937
Qy 5138 ACAGGCGATCGTGGTGTACGCTCGTGGTGTAGTGGTTCATTCAGCTCCGTTCCCAA 5197
Db 4938 ACAGGCGATCGTGGTGTACGCTCGTGGTGTAGTGGTTCATTCAGCTCCGTTCCCAA 4997
Qy 5198 CGATCAAGCGCAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGCT 5257
Db 4998 CGATCAAGCGCAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGCT 5057
Qy 5258 CTTCCGATCGTTGTGAGAAAGTAAGTTGGCGCAGTGTATCATCATGTTATGGCAGCA 5317
Db 5058 CTTCCGATCGTTGTGAGAAAGTAAGTTGGCGCAGTGTATCATCATGTTATGGCAGCA 5117
Qy 5318 CTGCATAATCTCTTACTGTCTAGCCATCGTGAAGATGCTTTTCTGTGACTGGTGAAGTAC 5377
Db 5118 CTGCATAATCTCTTACTGTCTAGCCATCGTGAAGATGCTTTTCTGTGACTGGTGAAGTAC 5177
Qy 5378 TCAACCAAGTCATTCTGAGAATAAGTGTATGCGCGCAGCCGAGTTGCTTCTCCCGCGGCTCA 5437
Db 5178 TCAACCAAGTCATTCTGAGANAAGTGTATGCGCGCAGCCGAGTTGCTTCTCCCGCGGCTCA 5237
Qy 5438 ATACGGGATAATAACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACCGT 5497
Db 5238 ATACGGGATAATAACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACCGT 5297
Qy 5498 TCTTCCGGGCGGAAACTCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCC 5557
Db 5298 TCTTCCGGGCGGAAACTCTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCC 5357
Qy 5558 ACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACCTTTTACCAGCGTTTCTGGGTGAGCA 5617
Db 5358 ACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACCTTTTACCAGCGTTTCTGGGTGAGCA 5417
Qy 5618 AAAACAGGAAGGCAAAATGCGCAAAAAAGGGAATAAGGCGCAGACGGAATTTGTAATA 5677
Db 5418 AAAACAGGAAGGCAAAATGCGCAAAAAAGGGAATAAGGCGCAGACGGAATTTGTAATA 5477
Qy 5678 CTCACTACTCTTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTTGCTCATGAGC 5737
Db 5478 CTAAATCTCTTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTTGCTCATGAGC 5537
Qy 5738 GGATACATATTTGAATGTATTTAGAAAAATTAACAATAGGGGTTCCGCGCACAATTTCCC 5797
Db 5538 GGATACATATTTGAATGTATTTAGAAAAATTAACAATAGGGGTTCCGCGCACAATTTCCC 5597
Qy 5798 CGAAAAAGTCCCACTGACGTC 5818
Db 5598 CGAAAAAGTCCCACTGACGTC 5618

Search completed: April 30, 2005, 10:34:04
Job time : 1097 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2005, 05:45:13 ; Search time 3019 Seconds
(without alignments)

11860.806 Million cell updates/sec

Title: US-10-715-665-6

Perfect score: 5882

Sequence: 1 tcgcgcgttcggtgatgac.....tatcacgaggcccttcgctc 5882

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 11284434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2654.4	45.1	10263	18	US-10-609-019-2
2	2419.2	41.1	6236	18	US-10-491-121-41
3	2397.8	40.8	6050	17	US-09-491-974-4
4	2397.8	40.8	6050	17	US-10-394-388A-4
5	2395.6	40.7	6148	18	US-10-491-121-24
6	2363	40.2	6258	18	US-10-491-121-34
7	2331	39.6	7685	19	US-10-763-976A-14
8	2321.4	39.5	7626	17	US-10-401-000-2
9	2311.8	39.3	7617	17	US-10-401-000-1
10	2290.6	38.9	6577	17	US-10-359-120-50
11	2283.4	38.8	6505	17	US-10-359-120-15
					Sequence 2, Appli
					Sequence 41, Appl
					Sequence 4, Appli
					Sequence 24, Appl
					Sequence 34, Appl
					Sequence 14, Appl
					Sequence 2, Appli
					Sequence 1, Appli
					Sequence 50, Appl
					Sequence 15, Appl

Sequence 34, Appl
Sequence 51, Appl
Sequence 157, App
Sequence 32, Appl
Sequence 100, App
Sequence 8, Appli
Sequence 158, App
Sequence 105, App
Sequence 106, App
Sequence 33, Appl
Sequence 122, App
Sequence 103, App
Sequence 84, Appl
Sequence 107, App
Sequence 16, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 115, App
Sequence 3, Appli
Sequence 4, Appli
Sequence 144, App
Sequence 130, App
Sequence 145, App
Sequence 128, App
Sequence 114, App
Sequence 108, App
Sequence 99, Appl
Sequence 67, Appl
Sequence 10, Appl
Sequence 98, Appl
Sequence 116, App
Sequence 101, App
Sequence 29, Appl
Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-10-609-019-2
; Sequence 2, Application US/10609019
; Publication No. US20040197910A1
; GENERAL INFORMATION:
; APPLICANT: Cooper, Richard K.
; APPLICANT: Cadd, Gary G.
; APPLICANT: Fioretti, William C.
; APPLICANT: DeBoer, Kenneth F.
; TITLE OF INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based
; FILE OF INVENTION: Vector
; FILE REFERENCE: 51687-0101 (51687-287015)
; CURRENT APPLICATION NUMBER: US/10/609,019
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/392,415
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/441,392
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,377
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,502
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,405
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,447
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US 60/441,381
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 10263
; TYPE: DNA
; ORGANISM: Artificial Sequence

Dbb 6096 CATCGCGTGAACCTTCCCTCCGAGCGCCCGTAAATGCAAGAAGAACCATGGCTGGGA 6155
Qy 2297 ACCCTTTACGATTTTGACAGGGCTGGGGCCCTATCAGTTATGCCAAACGAGCGGCC 2356
Dbb 6156 GGCCTCACGAGCGCTGTACCCCGGAGGGCTGCTGAAGGCGAGATCCACAAGGC 6215
Qy 2357 CGACAGCGCCCTACTGCTGGCACTACC-----CCCCAA 2391
Dbb 6216 CCTGAAGCTGAAGAGCGCGGCCACTACTCTGGTGGAGTTCAAGTCCATCTACATGGCCAA 6275
Qy 2392 AACCTTCGGGTATTTGTCGCGGAGAGTGTGTGTCGGGTATTTGCTTCACTCCC- 2450
Dbb 6276 GAAGCCGTCGAGTCCCGGCTACTACTACGTGGACTCCAAGCTGGACATCACTCCCA 6335
Qy 2451 -----AGCCCGTGTGTGGGAACGACCGACAGTCCGGGCGCCCACTACA 2499
Dbb 6336 CAACGAGGACTACACCATCTGGAGCAGTAGAGGCGACCGAGGCGCCACCACTGTT 6395
Qy 2500 GCTGGGTGAAATGA-----TACGGACGCTTTCGT 2530
Dbb 6396 CTTGTAGCGCGCGACTCTAGATCATATCAGCCATACCACATTTTGTAGAGGTTTACT 6455
Qy 2531 CCTTAACAATACGAGGCGCGTGGCAATTTGTTGGTTGTACCTGGATGAATCAAC 2590
Dbb 6456 TGGTTTAAAAAACCTCCCAACCTCCCTGAACTGAAACATATAAATGAATGCAATGT 6515
Qy 2591 TGGATTCAACAAAGTGTGCGAGCGCTCCTGTGTCTATCGGAGGCGGGCAACAAC 2650
Dbb 6516 TGTGTAACTGTTATTCGAGCTTAATATGTTACAAATAAAGCAATAGCATCAAAA 6575
Qy 2651 CCT-----GCACTGCCCCCACTGATTC 2672
Dbb 6576 TTTTCAAAATAAAGCAATTTTTCACGTCAATCTAGTTGTGGCCCGGCTGCAGGAATTC 6635
Qy 2673 TTCGGAAGCATCGGAGCGCAATCTCTCGTGGCTCCGGTCCCTGGATCAACCC 2732
Dbb 6636 GATATCAAGCTTATCGGATACCGCTGACCTCGAGGGGGGGCCCGTACCCCAATTCGCCCTA 6695
Qy 2733 AGTGCTGCTGCACTACCCG--TATAGGCTTTGGCATTATCTGTACCATCAATACA 2790
Dbb 6696 TAGTGAGTGTATACGCGGCTCACTGGCGCTGTTTTACAGGTGTGCTGGGAAA 6755
Qy 2791 CCATATTTAAATCAGGATGACGTGGGAGGGTTCGAACACAGGCTGGAAGCTCCCTGCA 2850
Dbb 6756 CCTGGGCTTACCAACTTAATCGCTTGCAGCACATCCCTTTTGCACGCTGGCGTAA 6815
Qy 2851 ACTGAGCGGGGGAACGTTGCTGATCTGGAAGATAGGACAGTCC----- 2897
Dbb 6816 TAGCAAGAGGCGCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATG 6875
Qy 2898 GAGATCGATATGAGAACATCAGATCAGGATTCCTAGGACCCCTGCTCGTTACAGGG 2957
Dbb 6876 GAAATGTAGCGTTAATATTTTGTGTTAAATTCGGTTAAATTTTGTGTTAAATCAGCTCA 6935
Qy 2958 GGGTTTTTCTGTGCAAGAATCCTCACATACCG----- 2993
Dbb 6936 TTTTAAACCAATAGGCGGAAATCGGAAATCCCTTATAAATCAAAAGATAGACGAG 6995
Qy 2994 -----CAGAGCTAGACTCGTGGTGGCACTTCTCTC 3023
Dbb 6996 ATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTTGACTCC 7055
Qy 3024 AATTTCTAGGGGATCTCCCGTGTCTTGGCCA----- 3058
Dbb 7056 AACGTAAAGGGGAAACCGTCTATCAGGGCGATGGGCCACTACTCCGGGATCATATG 7115
Qy 3059 ----- 3058
Dbb 7116 ACAAGATGTATCCACTTAACTTAATGATTTTACCAAAATCAATTAGGGGATTCATCA 7175
Qy 3059 ----- 3058
Dbb 7176 GTGCTCAGGGTCAACGAGAAATAACATTCCTCGTCAGGAAAGCTTATGATGATGATGTGCTT 7235

Qy 3059 -----AAATTCGAGTCCCAACCTCCAATCACTCACCACCTC 3097
Dbb 7236 AAAAACTTACTCAATGGCTGGTTATGCATATCGCAATACATGCAAAAAACCTAAAGAGC 7295
Qy 3098 CTGTCCTCCAAATTTGCTCGGTATCGCTGGATGTGTCTGCGGGCTTT----- 3145
Dbb 7296 TTGCCGATAAAAAAGGCCAATTTATTTGCTATTTACCGCGCTTTTATTGAGCTTGAAG 7355
Qy 3146 -----TATCATATTCCTTCTTCACTCTCTGCTGCTATGCTCATCTTCTT 3187
Dbb 7356 ATAAATAAATAGATAGGTTTATTTGAAGCTAAATCTTCTTATCGTAAAAAATGCCCT 7415
Qy 3188 ATTGGTTCTTCTGGAATTAACAAGGTATGTGCCCG----- 3222
Dbb 7416 CTTGGGTATCAAGAGGGTCATTATATTTCCGCGAATAACATCATTTGGTGACAAATAA 7475
Qy 3223 ----- 3222
Dbb 7476 CTAAGCACTTGTCTCTGTTTACTCCCTCAGCTTGAGGGGTTAAACATGAAGGTCATCGA 7535
Qy 3223 -----T 3223
Dbb 7536 TAGCAGGATAATAATACGTAAACCGCTAAACCAATAATCAAAATCCAGCCATCCCAAT 7595
Qy 3224 TTGTCCTCTAAATTCAGGATCAACAACACAGTACGGGACCAATGCAAAACCTGCACGAC 3283
Dbb 7596 TGGTAGTGAATGATTAATAACAGCAACAGTAATGGCCAAATAACACCGGTGTGATTT 7655
Qy 3284 TCCTGCTCAAGGCACTCTATGTTTCCCTCATGTGCTG----- 3322
Dbb 7656 GGTAAAGGCTCACCAATAATCCCTGTAAAGCACCTTGTGATGACTCTTTTGGATAGA 7715
Qy 3323 -----TACAAAACCTACGGATGGAATTTGCACCTG 3352
Dbb 7716 CATCACTCCCTGTAATGCAAGTAAAGCGATCCCAACAGCAATATAAATAAAGCAGG 7775
Qy 3353 TATTCCTCATCCCTCGCTCGCTGGGCTTTCGAAAATACCTATGGGAGTGGGCTCAGTCCG 3412
Dbb 7776 GAAACTAACCAACCTTCAGATATAAACGCTAAAGAGCAATGCACTACTATCTGCAAT 7835
Qy 3413 TTTCTCTGGCTCAGTTTACTAGTGCCATTTGTTCAAGTGTGTTAGGGCTTTCCGCCAC 3472
Dbb 7836 AAATCCGAGCAGTACTGCGGTTTTTCGCCCAATTTAGTGGCTATTCTCTCGCACAAAG 7895
Qy 3473 TGTGTTGGCTTTCAGCTATA---TGATGATGTGTTATTTGGGGGCCAAG-----TCTGT 3522
Dbb 7896 GCTTGAATACTGAGTGTAAAGACCAAGACCCGTAATGAAAAGCCAAACCATCATGCTAT 7955
Qy 3523 ACAGCATCGTGAAGTCCCTTTATACCGCTGTACCAATTTTCTTCTCTGGGTATACA 3582
Dbb 7956 TCATCATCAGATTTCTGTAAATAGCACACACGCTGCTGGATTCATCAATGCGCTGA 8015
Qy 3583 TTTAAGAATTCAGACTCGAGCAAGTCTAGAAGGCGCGCAAGATATCAAGATCCACTA 3642
Dbb 8016 AATAATAACAAATAATGCGATCGTTAAATAGTGATGATATACCGATCAGCTTTTGTTC 8075
Qy 3643 CGGTTAGAGCTCGCTGATCAGCTCGAC-----TGTCCTTCTAGTTGCCAGCCATC 3695
Dbb 8076 CTTTAGTAGGGTTAATGCGCGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGA 8135
Qy 3696 TGTGTTTTCCTCCCTCCCGCTGCTTCTTACCTCTGGAAGGTGCGACCTCCCACTGCTCT 3755
Dbb 8136 AATTGTTATCCGCTCAATTTCCACACATACAGAGCGGAGCATAAAGTGTAAAGCC 8195
Qy 3756 TTTCTTAATAATAGGAAATTCGATCGCAATTTCTGAGTAGGTGTCTTATTTCTGGG 3815
Dbb 8196 TGGGGTGCCTAATAGTAGAGCTAACTCAATTAATTTGCGTTGGCTCACTGCCCGCTTC 8255
Qy 3816 GGGTGGG---TGGGGCAGCAGCAGGGGGAGGATTTGGGAAGCAATAGCAGGATGC 3872
Dbb 8256 CAGTCGGGAAACCTGTGTCGTCAGCTGCTTAAATGAATGGCCACCGCGGGGAGAGGC 8315


```
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 6236
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Lassa (codon optimized)
; us-10-491-121-41

Query Match      41.1%; Score 2419,2; DB 18; Length 6236;
Best Local Similarity 70.9%; Pred. No. 0;
Matches 3491; Conservative 0; Mismatches 1298; Indels 133; Gaps 16;

QY 1 TC CGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
DB 1 TC CGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGGTG 120
DB 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAGCCCGTCAAGGCGCGTCAAGCGGGTG 120
QY 121 TTGCGGGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATGTTACTCAGAGTGC 180
DB 121 TTGCGGGGTGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATGTTACTCAGAGTGC 180
QY 181 ACCATATGAAGCTTTTTCAGAAAGCCTTAGGCTCCAAAGAGCTCCTCACTACTTCTGG 240
DB 181 ACCATATGCGG-----TGTAATACCGACA 207
QY 241 AATAGCTCAGAGCGGAGGGGCTCGGCCCTCTGCTATAATAAAAAAATAGTCAGCCA 300
DB 208 GATCGCTAAGGAGAAAAATACCGCATCAG-----235
QY 301 TGGGGCGGAGAAATGGCGGAACTGGGCGGAGGGAATATTGGCTATTGGCCATTGCAT 360
DB 236 -----ATTGGCTATTGGCCATTGCAT 256
QY 361 ACCTGTTATCTATATCATATATGATGATTAATTGGCTCATGTCCAAATATACCGCCA 420
DB 257 ACCTGTTATCTATATCATATATGATGATTAATTGGCTCATGTCCAAATATACCGCCA 316
QY 421 TGTGACATTGATTTAGTACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTCA 480
DB 317 TGTGACATTGATTTAGTACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTCA 376
QY 481 AGCCCATATATGAGTTCCGGTTACATACCTTACGTTAAATGGCCCGCTGCTGACCG 540
DB 377 AGCCCATATATGAGTTCCGGTTACATACCTTACGTTAAATGGCCCGCTGCTGACCG 436
QY 541 CCCAAGCAGCCCGCCCAATTGAGCTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600
DB 437 CCCAAGCAGCCCGCCCAATTGAGCTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 496
QY 601 GGGACTTTCCATTGACGTCAATCGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 660
DB 497 GGGACTTTCCATTGACGTCAATCGGTGGAGTATTTACGGTAAACTGCCCACTTTGGCAGTA 556
QY 661 CATCAGTGTATCATATGCGAGTCCGCCCTTATTGACGTCAATGACGGTAATGGCCC 720
DB 557 CATCAGTGTATCATATGCGAGTCCGCCCTTATTGACGTCAATGACGGTAATGGCCC 616
QY 721 GCCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
DB 617 GCCTGGCATTTATGCCAGTACATGACCTTATGGGACTTTTCTACTTTGGCAGTACATCTAC 676
QY 781 GTATTAGTCATCGCTATTACATGGTGTGCGGTTTGGCAGTACCAATGGCGGTGA 840
DB -----
```

1915 ATAGCTGACAGATAACAGACTGTTCTCTTCCATGGGTCTTTTCTCGAGTCAACCGTCGTC 1974
1817 ATAGCTGACAGATAACAGACTGTTCTCTTCCATGGGTCTTTTCTCGAGTCAACCGTCGTC 1876
1975 GACGAATTCAAGCAATCATGGATG-CAATCAAGAGAGGCTCTGCTGTGCTGCTGCTG 2033
1877 GACACGTGTGATCAGATATCGCGCGCTCTAGAGATATCGCGCATCGGCGCAGATCGT 1936
2034 TGTGGAGCAGTCTTCGTTTCGCCACAGGCTAGGAAACCCACG-TCACCGGGGGAAGTGC 2092
1937 GACCTTCTCCAGGAGTGCCCATGTGATCGAGGAGTGATGAACATCGTGTGATCGC 1996
2093 CGGCCACACTGTGCTGTGGATTTGTAGCTCTCTCGCACGAGGCGCCAGAGAAAGTCCA 2152
1997 CTTGAGCGTGTGGCGCTGTGAAGGCGCTGTACAACTTCGCCACCTCGCGCGCTGTGGG 2056
2153 GCTGATCAACACCAACCGCAGTTGGCACCTCAATAGCAGCGCCCTGAACGTGAATGATAG 2212
2057 CTTGGTGACCTTCTCTGCTGTGTGGCGAGGAGCTGCACACCGCTGTACAAAGGCGT 2116
2213 CTTCAACACCGGTGTGTGGCAGGCTTTTCTATCACCAAGTTCAACTCTTTTCAGGCTG 2272
2117 GTACGAGCTGCAGACCTCGAGCTGAACATGGAGACCTTGAAATGACCATGCCCTGAG 2176
2273 TCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTTGTACAGGCTGTGGGCGCCAT 2332
2177 CTGCACCAAGAAACACAGCACCATACATCATGTGTGGGCAACGACCGGCTGTGAGCT 2236
2333 CAGTTATGCCAAGGAGGCGCCCGACGAGCGCCCTACTGTGTCGACTACCCCCCAA 2392
2237 AACCTTGACCAACACAGCATCATCAACCAAGTTCTGCAACTGAGCGAGCGCCACAA 2296
2393 ACCTTCGGGTATTGTCCCGC---GAAGAGTGTGTGTGTCGGTATATTGCTTCACTCC 2449
2297 GAAGAACCTGTAGACCAACGCCCTGTAGCATCATCAGACACTTCCACTGAGCATCC 2356
2450 CAGCCCCGTGTGTGGGACGACGAGTCAGGTGGGCGGCCCACTCAAGCTGTGGGTGA 2509
2357 CAACCTCAACAGTACGAGGCCATGAGCTGCGACTTCAACGGCGGCAAGATCAGCGTGCA 2416
2510 AATGATAGCGAGCTTCTGTCCTTAAACAATACAGGCCACCGCTGGGCAATTGTTGCG 2569
2417 GTACAACCTGAGCCAAGCTACCGCGGCGACGCGCCAAACCACTGCGGCACCGTGCCAA 2476
2570 TTGTACTGTGATGAATCAACTGGAATTCAACAAAGTGTGGGAGCGCCCTCTGTGTCTAT 2629
2477 CGGCGTGTGCAGACCTTATGAGGATGGCTGGGCGGCGAGCTACATCGCCCTGGACAG 2536
2630 CGGAGGGCGGGCAACAAACCTGTCACTGCCCCACTGATGTTGTCGGAAGCATCCGGA 2689
2537 CGGCAAGGGCACTGGGACTGTCATCATGACCAAGCTACCAAGTACCTGATCATCCAGAAC 2596
2690 CGCCACATACTCTCGTGTGGGCTCCCGTCCCTGGATCACACGAGTGCCTGTGCTGACTA 2749
2597 CACTTGGGAGGACCACTGCGAGTTACGAGGCGCCCGCCCATCGGTACTTGGGCGCTGT 2656
2750 CCGTATAGGCTTTGGCATATTCTTTGTACCACTCAACTACCATATTTAAATACAGGAT 2809
2657 GAGCCAGAGGACCAGGACATCTACATCAGCAGAGGCTGCTGGGCACCTTCACTCTGAC 2716
2810 GTACGTGGAGGGGTGCAACACAGGCTGGAGTGTGCTGCAACTGGAACGCGGGGCGAACG 2869
2717 CTTGAGCGCAGCGAGGGCAAGACACACCGCGCGCTACTGCTGACCAAGTGTGCT 2776
2870 TTGCGATCTGGAAGATAGGACAGGTCGAGATTCGATATGGAGAACATCACATCAGGATT 2929
2777 GATCGAGGCGGAGCTGAATGTCTTCGGCAACACCGCGGTGGCCAGTGCACGAGAACGA 2836
2930 CTTAGGACCCCTGCTGCTGTATTACAGCGGGGTTTTTTCTTTGTTGACAAGAAATCTCTCAAT 2989
2837 CGACGAGGAGTTCTCGACATGCTGAGGCTGTTTCGACTTCAACAAACGAGGCCATCCAGAG 2896

2990 ACCGACAGTCTAGACTCGTGTGGACTTCTCTCAATTTTCTAG-GGGGATCTCCCGTGT 3048
2897 GCTGAGGCGGAGGCCAGATGAGCATCAGCTGATCAACAAGGCGGTGAACGCGCTGAT 2956
3049 GTCTTGGCCAAAATTCGAGTCCCAACCTCAATCAGCTCAACCACTCTGTCTCTCCAA 3108
2957 CAACGACGAGTATCATGAAGAACACCTGAGGACATCATGGGCATCCCTACTACTGCAA 3016
3109 TTTGTCTCTGGTTATCGCTGGATGTGTCTCGGC-----GTTTTATCATATTTCTCTTCAT 3163
3017 CTACAGCAAGTACTGGTACTGAACACACACCAACCAACCGGACGAGCTGCCAAGTG 3076
3164 CTTGCTGTATCGCTCATCTTCTTATTG-GTTCTTCTGATTTATCAAGTATGTTGCCCG 3222
3077 CTGGCTGTGACCAACGCGCAGCTACTGAAACGAGACCCACTTTCAGCGACGACATCGAGCA 3136
3223 TTTGTCTCTTAATTCAGGATCAACAAACACAGTACGGGACCATGCACAACTCGACCA 3282
3137 GCAGGCGACATGATCACGAGATGCTGCAGAGGAGTACATGGAGGAGCGAGGCAA 3196
3283 CTCCTGCTCAAGGCAACTCTATGTTTCC---CTCATGTTGTGTATACAAACCTACGAGTG 3339
3197 GACCTGAAACACGTGGATCCAGATCTGTGTGCTTCTAGTTGCCAGCCATCTGTTGTTT 3256
3340 GAAATTGCACCTGTATTTCCATCCCATCGTCTGGCTTTTCGCAAAATACCTATGGAGT 3399
3257 GCCCTTCCCGCTGCTCTTTCAGCTGGAAGGTGCCACTCCCACTGTCTTCTCTTAAT 3316
3400 GGGCTCTAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCATTGTTTCAAGTGTTCGTAG 3459
3317 AAAATGAGAAATTGCATGCAATGCTCTGAGTAGGTGTCATTCTATTCTGGGGGTGGG 3376
3460 GGGCTTCCCGCACTGTTGGCTTTTCACTATATGATGATGTTGTTATGGGGCCAAAGTC 3519
3377 TGGGCGAGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGATGCTGGGATGCGG 3436
3520 TGACAGCATCTGTGAGTCCCTTTATACCGCTTTTACCAATTTTCTTTTCTCTCTGGTAT 3579
3437 TGGGCTCTATGGGTACCCAGGTGCTGAAGAAATTGACCCCGGTTCTCTCTGGGCCAAGA 3496
3580 ACATTTAAGAAATTCAAGCTCGAGCAAGTCTAGAAAGGCGCGCAAGATATCAAGATCCA 3639
3497 AGCAGGCACATCCCTTCTC-TGTGACACACCTGTCCAGCGCCCTGTGTTCTTAGTTCCA 3555
3640 CTACGCTTGTAGCTGCTGATCAGCTCGACTGTGCTTCTAGTTGTCAGGCCATCTGTT 3699
3556 GCCCACTCATAGGACACTCATAGCTCAGGAGGGCTCCGCTTCAATCCACCCGCTAAA 3615
3700 GTTTGCCCTCCCGCTGCTTCTTGACC-----CTGGAAGTGCCTACTCCACTGTC 3753
3616 GTACTTGGAGCGGTCTCTCCCTCCCTCATAGCCCAACCAACCAACCTAGCCCTCCAAG 3675
3754 CTTTCTTAATAAATGAGGAAATTGCATCGCAATTTGCTGAGTAGTGTGTCATTCTATTCTG 3813
3676 GTGGGAGAAATTAAGCAAGATAGGCTATTAGTGCAGAGGAGAGAAATGCTCTCAA 3735
3814 GGGGTGGGGTGGGCGAGCAGCAAGGGGAGGATTTGGGAAGACAATAGCAGGCAATGCT 3873
3736 CATGTGAGGAATTAATGAGAGAAATCATAGAATTTTAAGGCCATGATTTTAAGGCCATCAT 3795
3874 GG-GGAGCTCTTCCGCTTCTCGCTCAGTACTCGCTGCGCTCGGTCGTTTCGGCTCGCGC 3932
3796 GGGCTTAATCTTCCGCTTCTCGCTCAGTACTCGCTGCGCTCGGTCGTTTCGGCTCGCGC 3855
3933 GAGCGGTATCAGTCACTCAAAAGGGGTAATACGGTTATCCACAGAAATCAGGGGATAACG 3992
3856 GAGCGGTATCAGTCACTCAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACG 3915
3993 CAGGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACGTAATAAGGCCCGCT 4052
3916 CAGGAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACGTAATAAGGCCCGCT 3975
4053 TGCTGGCGTTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAATTCAGCGCTCAA 4112

Db 3976 TGCTGGCGTTTTCATAGCTCGGCCCTCAGCAGCATCAAAAAATCGAGCTCAA 4035
QY 4113 GTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCTCGGAAGCT 4172
Db 4036 GTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCTCGGAAGCT 4095
QY 4173 CCTCGTGGCTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCC 4232
Db 4096 CCTCGTGGCTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCC 4155
QY 4233 CTTTGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTGGGTAGG 4292
Db 4156 CTTTGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTGGGTAGG 4215
QY 4293 TCGTTCGCTCAAGCTGGGCTGTGTCAGAACCCCGTTTACGCGCGACCGCTGCGCT 4352
Db 4216 TCGTTCGCTCAAGCTGGGCTGTGTCAGAACCCCGTTTACGCGCGACCGCTGCGCT 4275
QY 4353 TATCCGGTAACTATCGTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAG 4412
Db 4276 TATCCGGTAACTATCGTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAG 4335
QY 4413 CAGCCACTGTAACAGGATAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGA 4472
Db 4336 CAGCCACTGTAACAGGATAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGA 4395
QY 4473 AGTGTGGCTTAACCTACTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGA 4532
Db 4396 AGTGTGGCTTAACCTACTACGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGA 4455
QY 4533 AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCAACCGCTG 4592
Db 4456 AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCAACCGCTG 4515
QY 4593 GTAGCGGTGGTTTTTTTGTGTTGCAAGCAGAGATTAACGCGAGAAAAAAGAGTCTCAAG 4652
Db 4516 GTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTAACGCGAGAAAAAAGAGTCTCAAG 4575
QY 4653 AAGATCCTTTGATCTTTTCTACGGGCTCTCAGCTCAGTGGAGCAAACTCAGGTTAG 4712
Db 4576 AAGATCCTTTGATCTTTTCTACGGGCTCTCAGCTCAGTGGAGCAAACTCAGGTTAG 4635
QY 4713 GGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAAT 4772
Db 4636 GGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAATAAT 4695
QY 4773 GAAGTTTTAAATCAATCTAAAGATATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 4832
Db 4696 GAAGTTTTAAATCAATCTAAAGATATATAGTAAACTTTGGTCTGACAGTTACCAATGCT 4755
QY 4833 TAATCAGTCAGGCACCTATCTCAGCGCATCTGTCTATTTCGTTTCATCCATAGTTGCTGAC 4892
Db 4756 TAATCAGTCAGGCACCTATCTCAGCGCATCTGTCTATTTCGTTTCATCCATAGTTGCTGAC 4815
QY 4893 TC 4894
Db 4816 TC 4817

RESULT 3

US-09-491-974-4
; Sequence 4, Application US/09491974
; Patent No. US200201148181
; GENERAL INFORMATION:
; APPLICANT: U.S. Medical Research Institute of Infectious Diseases
; APPLICANT: Schmaljohn, Connie S.
; APPLICANT: Hooper, J. W.
; TITLE OF INVENTION: DNA Vaccines Against Hantavirus Infections
; FILE REFERENCE: Army-142
; CURRENT APPLICATION NUMBER: US/09/491,974
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,680

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Microsoft Word, Office 97, IBM compatible
; SEQ ID NO 4
; LENGTH: 6050
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: /note="S genome segment from SBOV, strain SR-11, subcloned into D

Query Match 40.8%; Score 2397.8; DB 9; Length 6050;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 3456; Conservative 0; Mismatches 1097; Indels 240; Gaps 16;
QY 337 ATTATTGGCTATTGGCCATTGTCATAGCTTGTATCTATATCATATATATGACATTATTTATTTATTT 396
Db 1263 AATATTGGCTATTGGCCATTGTCATAGCTTGTATCTATATCATATATATGACATTATTTATTTATTT 1322
QY 397 GGCTCATGTCCCAATATGACCGCATGTTGACATTTGATTATTTGACTAGTTTATTAATAGTAA 456
Db 1323 GGCTCATGTCCCAATATGACCGCATGTTGACATTTGATTATTTGACTAGTTTATTAATAGTAA 1382
QY 457 TCAATTACCGGCTCATTTAGTTTCATAGCCCATATATAGGAGTTCGCGTTACATAACTTACG 516
Db 1383 TCAATTACCGGCTCATTTAGTTTCATAGCCCATATATAGGAGTTCGCGTTACATAACTTACG 1442
QY 517 GTAAATGGCCCGCTGGCTGACGGCCCAACGACCCCGCCCATTTGACGTCAATAATGACG 576
Db 1443 GTAAATGGCCCGCTGGCTG-TGACGGCCCAACGACCCCGCCCATTTGACGTCAATAATGACG 1501
QY 577 TATGTTCCCATAGTAACGCAATAGGGAATTTCCATTGACGTCAATGGGTGGAGTATTTA 636
Db 1502 TATGTTCCCATAGTAACGCAATAGGGAATTTCCATTGACGTCAATGGGTGGAGTATTTA 1561
QY 637 CGGTAAATGCCACTTGGCAGTACATCAAGTGTATCATATGCAAGTCC-GCCCCCTAT 695
Db 1562 CGGTAAATGCCACTTGGCAGTACATCAAGTGTATCATATGCAAGTCCGGCCCCCTAT 1621
QY 696 TCAGCTCAATGAGCGTAAATGGCCCGCTGGCATTTATGCCAGTACATGACCTTAGCGGA 755
Db 1622 TCAGCTCAATGAGCGTAAATGGCCCGCTGGCATTTATGCCAGTACATGACCTTAGCGGA 1681
QY 756 CTTTCTCTACTTGGCAGTACATCTACGTATTAGTCTCGCTATTACCATGCTGATGCGGTT 815
Db 1682 CTTTCTCTACTTGGCAGTACATCTACGTATTAGTCTCGCTATTACCATGCTGATGCGGTT 1741
QY 816 TTGGCAGTACACCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTCCA 875
Db 1742 TTGGCAGTACACCAATGGCGGTGGATAGCGTTTGACTCAGCGGGATTTCCAAAGTCTCCA 1801
QY 876 CCCCATTGAGCTCAATGGGAGTTTGTGTCACCAAAATCAACGGGACTTTTCCAAATG 935
Db 1802 CCCCATTGAGCTCAATGGGAGTTTGTGTCACCAAAATCAACGGGACTTTTCCAAATG 1861
QY 936 TCGTAATAACCCCGCTTTCAGCAATGGCGGTAGCGGTAGCGGTAGCGGTAGCGGTAGCGGTAG 995
Db 1862 TCGTAATAACCCCGCTTTCAGCAATGGCGGTAGCGGTAGCGGTAGCGGTAGCGGTAGCGGTAG 1921
QY 996 TATAAGCAGAGCTCGTTTATAGTGAACCGTTCAGATCGCTCGGAGACGCGCATCCACGCTGTTT 1055
Db 1922 TATAAGCAGAGCTCGTTTATAGTGAACCGTTCAGATCGCTCGGAGACGCGCATCCACGCTGTTT 1981
QY 1056 TGACCTCCATAGAGACACCGGACCGGATCCAGCTCCGCGGCGCGGAGCGGTGCAATTGG 1115
Db 1982 TGACCTCCATAGAGACACCGGACCGGATCCAGCTCCGCGGCGCGGAGCGGTGCAATTGG 2041
QY 1116 AACGGCGATTCCCGCTGCCAAGAGTACGTAAGTACCGCTATAGACTCTATAGGCACAC 1175
Db 2042 AACGGCGATTCCCGCTGCCAAGAGTACGTAAGTACCGCTATAGACTCTATAGGCACAC 2101
QY 1176 CCCTTTGGCTCTTATGCAATGCTATGTTTGGCTTGGGCGCTATACACCCCGCTCC 1235

Db 2102 CCCTTTGGCTCTATGTCATGCTATACGTGTTTTTGGCTTGGGCGCTATACACCCCGCTCC 2161
QY 1236 TTATGCTATAGGTGATGGTATAGCTTTAGCTATAGGTGTTGGTATTATGACCAATTTATGAC 1295
Db 2162 TTATGCTATAGGTGATGGTATAGCTTTAGCTATAGGTGTTGGTATTATGACCAATTTATGAC 2221
QY 1296 CACTCCCTTATTTGGTGACCATATTTTCCATTTACTTAATCCATTAACATGGCTCTTTGGCCACA 1355
Db 2222 CACTCCCTTATTTGGTGACCATATTTTCCATTTACTTAATCCATTAACATGGCTCTTTGGCCACA 2281
QY 1356 ACTATCTCTATTTGGCTATATGCAATPACTCTGTCTTTAGAGACTGACACGGACTCTGTGA 1415
Db 2282 ACTATCTCTATTTGGCTATATGCAATPACTCTGTCTTTAGAGACTGACACGGACTCTGTGA 2341
QY 1416 TTTTACAGGATGGGTG - CCAATTTATTTTAAATTTCAAAATTCACATATACAAACACGGCGTCC 1474
Db 2342 TTTTACAGGATGGGTGCGCAATTTATTTTAAATTTCAAAATTCACATATACAAACACGGCGTCC 2401
QY 1475 CCGGTGCCCGCAGTTTATTTAAACATAGCTGGGATCTCC---GACATCTCGGGTAGCT 1531
Db 2402 CCGGTGCCCGCAGTTTATTTAAACATAGCTGGGATCTCCACGCGAAATCTCGGGTAGCT 2461
QY 1532 GTTCCGGACATGGGCTCTTCTCCGGTAGCGGGAGCTTCCACATCCGAGCCCTGGTCC 1591
Db 2462 GTTCCGGACATGGGCTCTTCTCCGGTAGCGGGAGCTTCCACATCCGAGCCCTGGTCC 2521
QY 1592 ATCCGTCCAGCGGCTCATGGTCCGTCCGAGCTCTTTGTCTCTTAACAGTGGAGGCGACAG 1651
Db 2522 ATGCCCTCCAGCGGCTCATGGTCCGTCCGAGCTCTTTGTCTCTTAACAGTGGAGGCGACAG 2581
QY 1652 TTAGGCACAGCACAAATGCCACACACAGTGTGCCGACAGGCGGTGGCGGTAGGGT 1711
Db 2582 TTAGGCACAGCACAAATGCCACACACAGTGTGCCGACAGGCGGTGGCGGTAGGGT 2641
QY 1712 ATGTGTCTCAAAATGAGCTCGGAGATTGGGCTCGACCTGGACGCGAGATGGAGACTTAA 1771
Db 2642 ATGTGTCTCAAAATGAGCTCGGAGATTGGGCTCGACCTGGACGCGAGATGGAGACTTAA 2701
QY 1772 GGCAGCGGCAGAAAGATGCGAGGAGCTGAGTTGTTGTATTCTGATTAAGATGTCAGAGT 1831
Db 2702 GGCAGCGGCAGAAAGATGCGAGGAGCTGAGTTGTTGTATTCTGATTAAGATGTCAGAGT 2761
QY 1832 AACTCCGTTGGGTGCTGCTTAACGTTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGC 1891
Db 2762 AACTCCGTTGGGTGCTGCTTAACGTTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGC 2821
QY 1892 TGCAGCGCGCGCCACACAGACATAATAGCTGCAGACTTAAACAGACTGTTCTTTCCATGGG 1951
Db 2822 TGCAGCGCGCGCCACACAGACATAATAGCTGCAGACTTAAACAGACTGTTCTTTCCATGGG 2881
QY 1952 TCTTTTCTGAGTACCGTCTGTCGAGCAATTCAGCAATTCATGGATGCAATGAAGAGAG 2011
Db 2882 TCTTTTCTGAGTACCGTCTGTCGAGCAATTCAGCAATTCATGGATGCAATGAAGAGAG 2097
QY 2012 GCTCTGCTGTGTGCTGCTGTGTGGAGAGTCT---TCGTTTCCCGCCGCGCTAGCGA 2068
Db 2938 ACTCCCTTAAGAGCTACTACACTAAACAAGAAAATGGCAACTATGGAAGAAATCCAGAGA 2997
QY 2069 AACCCACGTCACGGGGGAAGTCCCGGCCACACTGTGTCTGGATTGTTAGGCTCTCTCGC 2128
Db 2998 GAAATCAGTGTCTACAGAGGGCGAGCTTGTGTAGCAGCGCCAGAGGTCAAGGATGCAGAA 3057
QY 2129 ACCAGGCGCAAGCAGAGAGCTGCAGCTGATCAACACCAACCGGAGTTGGCACTCTCAATAG 2188
Db 3058 AAGCAGTATGAGAAGGATCCTGATGACTTAAACAAGAGGGCACTGCATGATCGGAGAGT 3117
QY 2189 CACGGCCCTGAATGATGATAGCTCAACCGGCTGTGGCAGGCGCTTTCTATCA 2248
Db 3118 GTCCGAGCTTCAATACAAATTTGATGAATTTGAAGGCCAACTTCCGACAGAGTTG 3177
QY 2249 CCACAGTTCAACTCTTACAGGTGTCTCTGAGAGGCTAGCAGCTGCCGACCCCTTACCGA 2308
Db 3178 CAGCAGGGAAGAACATCCGCGCAGGACCGGGATCTCTACAGGGGTAGAGCCAGGTGATCAT 3237

QY 2309 TTTTGACAGGCGCTGGGCGCCTATCAGTTATGCCAAACGGAAGCGGCCCCCGAC----CAGC 2364
Db 3238 CTTAAGGAAGATACGCACTAAGCTACGGGAATACACTGGACCTGAATAGTCTTGACATT 3297
QY 2365 GCCCTACTGCTGGCACTACCCCCAAAACCTTGGCGTATTGTGCGCGGAAGAGTGTGT 2424
Db 3298 GATGAACCTACAGGACAGACAGCTGATTTGGCTGACCAATAATTTGTCTATCTGACATCTC 3357
QY 2425 GTGGTCCGG-----TATATTGCTTCACTCCAGCCCCCGTGTGGTGGGA 2468
Db 3358 GTGGTCCCGATCATCTTGAAGGCACTGTATACATGTTTAAACAACAGAGTAGGACACTTCA 3417
QY 2469 AGCAGCAGAGTCCGGCGCGCCCACTACAGCTGGGTGAAAATGATACGACGCTCTTC 2528
Db 3418 AAGGACAAACAGGGGATGAGGATCAGATTTCAAGGATGACAGCTCATATAGGATGTCAAT 3477
QY 2529 GTCCCTTAACAATPACAGGCGACCGCTGGGCAATTTGGTTGTATACCTGGATGAATCA 2588
Db 3478 GGAATCAGAAAGCCCAAAATCTGTATGTCTAATGCCAAACGCCCAATCCAGCATGAAG 3537
QY 2589 ACTGGATTCACAAAGTGTGGGAGCGCTCCTTGTGTATCGGAGGGCGGGCAACAAAC 2648
Db 3538 GCTGAAGAGATAAACACAGGAAGATTTCCGCACTGCAGTATGTGGACTATATCTGCACAG 3597
QY 2649 ACCCTGCACTGCCCACTGATTGTTCCGCAAGCATCCGAGCGCCACATCTCTCGGTGC 2708
Db 3598 ATAAAGGCAAGGAATATGTTAAGCCCTGTCTAGTGTAGTTGGGTCTTTGGCACTGGCA 3657
QY 2709 GGCTCCGGTCCCTGGATCACACCCAGGTGCCGTGCTGACTACCCGTATAGGCTTTGGCAT 2768
Db 3658 AAGACTGGACATCAGAAATTTGAAGATTTGGCTCGGTGCACCTTCGCAATTTATGGCGAG 3717
QY 2769 TATCCTTGTACC-----ATCAACTACACCATATTTTAAATCAGGATGTA 2812
Db 3718 TCTCTTATTTGCCGGAGTTTATCTGGAAATCTGTGAATCTGTACTATATCAGACAGACA 3777
QY 2813 CGTGGAGGGGTGCAACACAGCTGGAACTGCCCTGCACTGGACGCGGGCGGAAGTTG 2872
Db 3778 CAAGTGCATTTGCAAGGATGGAGCCAAAGGAATTTCAAGCCCTCAGGCAATTTCAAAG 3837
QY 2873 CGATCTGGAAGATAGGGACAGGTCCGAGATCGATATGGAGAACATCACATCAGGATTTCT 2932
Db 3838 GATGCTGAGTGTACACTAGTTGAAATATTTAGTACCATCATCAATATGGGTGTTGCT 3897
QY 2933 AGGACCCCTGCTGCTGTGTACAGCGGGGTTTTTCTTTGTGACAAGA----- 2978
Db 3898 GGGGCGCCTGATAGGTGTCCACCAACATGCTGTTGTTGTGGAGGATGGCTGAAATTAGGT 3957
QY 2979 ---ATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTCTAGG 3035
Db 3958 GCTTTCTTTCTATATCTCAGGATATGAGGAACAATCATGCTGGCTTCAAAAATCTGGG 4017
QY 3036 GGATCTCCCGTGTGTTGGCCAAAATTCGCACTCCCAACCTCCAATCACTCAC----- 3090
Db 4018 ACAGCTGATGAAGAGCTTCGAAGAAATCATCATTTCTATCAATCATACCTCAGACGACA 4077
QY 3091 CAACTCCTGCTCCTCAATTTGCTGTGTTATCGGTGAGATGTCTGCGCGGTTTTATCA 3150
Db 4078 CAATCAATGGGAATACAACTGGACAGAGATAAATTTGTTATGTTTATGTTGCTGGGGA 4137
QY 3151 TATTCTCTTCATCTGCTGCTATGCTCATCTTCTTATTGGTCTCTCGGATTTATCAAG 3210
Db 4138 AAGGAGGAGTGGACAACCTTTTCATCTCGGTGATGATGATGATGATGATGATGATGAT 4197
QY 3211 GTATGTTGCCGCTTGTCTCTTAATTTCCAGGATCAACAAACCAACAGTACGGGACCATGA 3270
Db 4198 GCTCAGATCTTGATTTGACCAAGAAAGTGAAGGAATCTCAAACAGGNACTTATGAATTA 4257
QY 3271 AAACCTGCACGACTCCTGCTCAAGGCACTCATGTTTTCCTCCCTCATGTTGCTGTACAAAC 3330
Db 4258 TAAGTACATAATAATAATCAATACTACTATAGTTTAAAGAAATACTAATCATTTATGTTA 4317

Query Match		40.8%;	Score 2397.8;	DB 17;	Length 6050;
Best Local Similarity		72.1%;	Pred. No. 0;		
Matches 3456;		Conservative	0;	Mismatches 1097;	Indels 240; Gaps 16;
QY	337	ATTATTGGCTATTGGCCATTGGCATAGTGTGTATCTATATCATATAATATGATACATTATATTT	396		
Db	1263	AATATTGGCTATTGGCCATTGGCATAGTGTGTATCTATATCATATAATATGATACATTATATTT	1322		
QY	397	GGCTCATGTCCAAATATGACCGCCATGTGACATTTGACATTTGACTAGTTATTAATAGTAA	456		
Db	1323	GGCTCATGTCCAAATATGACCGCCATGTGACATTTGACTAGTTATTAATAGTAA	1382		
QY	457	TCAAATTACGGGTTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACG	516		
Db	1383	TCAAATTACGGGTTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACG	1442		
QY	517	GTAAATGGCCCGCTGGCTGACCGCCCAACAGCCCCCGCCATTGACGTCAATAATGACG	576		
Db	1443	GTAAATGGCCCGCTCG-TGACCGCCCAACAGCCCCCGCCATTGACGTCAATAATGACG	1501		
QY	577	TATGTTCCCATAGTAACGCCAATAGGACTTTTCCATTGACGTCAATAGGGTGGAGTATTTA	636		
Db	1502	TATGTTCCCATAGTAACGCCAATAGGACTTTTCCATTGACGTCAATAGGGTGGAGTATTTA	1561		
QY	637	CGGTAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCC-GCCCCCTAT	695		
Db	1562	CGGTAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGGCCCCCTAT	1621		
QY	696	TGACGTCAATGACGGTAATAGGCCCGCTGGCATTTATGCCAGTACATGACACCTTTACGGGA	755		
Db	1622	TGACGTCAATGACGGTAATAGGCCCGCTGGCATTTATGCCAGTACATGACACCTTTACGGGA	1681		
QY	756	CTTTCCTACTTGGCAGTACATCTACGTATPAGTCATCGCTATTACCATGGTGTATGCGT	815		
Db	1682	CTTTCCTACTTGGCAGTACATCTACGTATPAGTCATCGCTATTACCATGGTGTATGCGT	1741		
QY	816	TTGGCAGTACACCAATGGCGTGGATAGCGGTTTGACTCAGCGGATTTCCAGTCTCCA	875		
Db	1742	TTGGCAGTACACCAATGGCGTGGATAGCGGTTTGACTCAGCGGATTTCCAGTCTCCA	1801		
QY	876	CCCATTTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG	935		
Db	1802	CCCATTTGACGTCAATGGGAGTTTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATG	1861		
QY	936	TCGTATAACCCCGCCCGTTGACGAAATPGGCGGTAGCGGTGATACGTTGGAGGTCTA	995		
Db	1862	TCGTATAACCCCGCCCGTTGACGAAATPGGCGGTAGCGGTGATACGTTGGAGGTCTA	1921		
QY	996	TATAAGCAGAGTCGTTTATGTGAACCGTCAGATCGCTCGGAGACGCATCCACGCTGTTT	1055		
Db	1922	TATAAGCAGAGTCGTTTATGTGAACCGTCAGATCGCTCGGAGACGCATCCACGCTGTTT	1981		
QY	1056	TGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCGCGCGCGGAAACGGTGCATTGG	1115		
Db	1982	TGACCTCCATAGAAGACACCGGACCGATCCAGCCTCCGCGCGCGGAAACGGTGCATTGG	2041		
QY	1116	AACCGGATTTCCCGTGCACGAAGTGAAGTGAAGTACGCGCTATAGACTCTATAGGCACAC	1175		
Db	2042	AACCGGATTTCCCGTGCACGAAGTGAAGTGAAGTACGCGCTATAGACTCTATAGGCACAC	2101		
QY	1176	CCCTTTGGCTCTTATGCATGCTATCTGTTTGGCTTGGGCTATACACCCCGCTCC	1235		
Db	2102	CCCTTTGGCTCTTATGCATGCTATCTGTTTGGCTTGGGCTATACACCCCGCTCC	2161		
QY	1236	TTATGCTATAGGTGATGGTATAGCTTAGCCCTATAGGTGTGGGTATTGACCATTTATGAC	1295		
Db	2162	TTATGCTATAGGTGATGGTATAGCTTAGCCCTATAGGTGTGGGTATTGACCATTTATGAC	2221		
QY	1296	CATCCCTCTATTGGTGACGATATTTCCATTAATTCATTAACATGCGTCTTTTGGCACA	1355		
Db	2222	CATCCCTCTATTGGTGACGATATTTCCATTAATTCATTAACATGCGTCTTTTGGCACA	2281		
QY	1356	ACTATCTCTATTGGCTATATGCCAATACTCTGTCTCTCAGAGACTGACCGGACTCTGTA	1415		

Db	2282	ACTATCTCTATTGGCTATATGCCAATACTCTGTCTCTCAGAGACTGACCGGACTCTGTA	2341		
QY	1416	TTTTTACAGATGGGT-CCATTATTATTTACAAATTCACATATACAAACCGCTCC	1474		
Db	2342	TTTTTACAGATGGGTGCCATTATTTATTTACAAATTCACATATACAAACCGCTCC	2401		
QY	1475	CCGTCGCCCGAGTTTTTTATTTAAACATAGCTGGGATCTCC---GACATCTCGGGTACGT	1531		
Db	2402	CCGTCGCCCGAGTTTTTTATTTAAACATAGCTGGGATCTCCACGCGAAATCTCGGGTACGT	2461		
QY	1532	GTTCCGGACATCGGCTCTTCTCCGTTAGCGGGAGCTTCCACATCCGAGCCCTGTGTCC	1591		
Db	2462	GTTCCGGACATCGGCTCTTCTCCGTTAGCGGGAGCTTCCACATCCGAGCCCTGTGTCC	2521		
QY	1592	ATCCGTCACAGCGCTCATGTCGTCGGCAGCTCTTGTCTCTAAACAGTGGAGCCAGAC	1651		
Db	2522	ATGCCTCCAGCGCTCATGTCGTCGGCAGCTCTTGTCTCTAAACAGTGGAGCCAGAC	2581		
QY	1652	TTAGGCACAGCAATGCCCCACACACAGTGTGCCGCACAAGCCGTGGCGGTAGGGT	1711		
Db	2582	TTAGGCACAGCAATGCCCCACACACAGTGTGCCGCACAAGCCGTGGCGGTAGGGT	2641		
QY	1712	ATGTGTCGAAATGAGCTCGGAGATTGGCTCGCACCTGGACGCAGATGGAAAGCTTAA	1771		
Db	2642	ATGTGTCGAAATGAGCTCGGAGATTGGCTCGCACCTGGACGCAGATGGAAAGCTTAA	2701		
QY	1772	GGCAGCGGACAGAAAGATGACGGCAGCTGAGTTGTGTATTCTGATAGAGTACAGGT	1831		
Db	2702	GGCAGCGGACAGAAAGATGACGGCAGCTGAGTTGTGTATTCTGATAGAGTACAGGT	2761		
QY	1832	AATCCCGTTGGGTGCTGTTAAACGTTGAGGGCAGTGTAGTCTGACGAGTACTCTGTC	1891		
Db	2762	AATCCCGTTGGGTGCTGTTAAACGTTGAGGGCAGTGTAGTCTGACGAGTACTCTGTC	2821		
QY	1892	TGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTCTCTTCCATGGG	1951		
Db	2822	TGCGCGCGCCACACAGACATAATAGCTGACAGACTTAAACAGACTGTCTCTTCCATGGG	2881		
QY	1952	TCCTTTCTCAGTCAACCGTCGTGCGAAGATTCAAGCAATCATGGATGCAATGAAGAGAG	2011		
Db	2882	TCCTTTCTCAGTCAACCGTCGTGCGAAGATTCAAGCAATCATGGATGCAATGAAGAGAG	2937		
QY	2012	GCTCTGCTGTGCTGCTGCTGTGGAGCAGTCT---TCGTTTCCCGCAGCGCTAGCGA	2068		
Db	2938	ACTCCCTCAAGAGCTACTACACTAAACAGAAAAATGGCAACTATGGAAGAAATCCAGAG	2997		
QY	2069	AACCCACGTCACCGGGGAAAGTCCCGCCACACTGTGTCTGGATTGTGTAGCCCTCTCGC	2128		
Db	2998	GAAATCAGTGTCTACGAGGGGCAGCTTGTGATAGCAGCGCCAGNAGGTCAAGGATGCAGAA	3057		
QY	2129	ACCAGCGCCAAAGCAGAACGTCAGCTGATCAACACCAACGCGAGTTGGCACCCTCAATAG	2188		
Db	3058	AAGCAGTATGAGAAAGGATCCTGATGACTTTAAACAGAGGGCACTGCATGATCGGGAGT	3117		
QY	2189	CAGGCGCTGAACTGCAATGATAGCCTCAACCGGCTGGTTGGCAGGGCTTTTCTATCA	2248		
Db	3118	GTGCGAGCTTCAATACAAATCAAAAAATTGATGAATTTGAAGCGCCAACTTGCCCGACAGATTG	3177		
QY	2249	CCACAGTTTCACTCTTCAGGGCTGCTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGA	2308		
Db	3178	CAGCAGGGAAGAACATCCGGGCAGGACCGGATCCTACAGGGGTAGAGCCAGGTGATCAT	3237		
QY	2309	TTTTTGACAGGGCTGGGGCCCTATCAGTTATGCAACGGAAGCGGCCCCGAC----CAGC	2364		
Db	3238	CTTAAGGAAGATCAGCACTAAGCTACGGGAATACACTGGACCTGAATAGTCTTTGACATT	3297		
QY	2365	GCCCTTACTGCTGGCACTACCCCCAAACCTTGGGTATGTGCCGCGAAGAGTGTGT	2424		
Db	3298	GATGAACCTACAGGACAGACAGCTGATTGGCTGACCATTAATTGCTATCTGACATCATTC	3357		
QY	2425	GTGGTCCGG-----TATATTGCTTCACTCCACGCCCCGTGGTGGGA	2468		

Db 3358 GTGGTCCCGATCATCTTTGAAGGCATGTATCATGTTAAACAACAGAGGTAGGCAGACTTCA 3417
Qy 2469 ACACCGACAGAGTTCGGCGCGCCACCTACAGCTGGGGTGAATAATGATACGGAGCTTTC 2528
Db 3418 AAGGACAAACAAGGGATGAGATCAGATTCAGAGATGACAGCTCATATGAGGATGTCAT 3477
Qy 2529 GTCCCTTAACAATACCAAGGCCACCGCTGGGCAATTTGGTTTCGGTTGTACCTCGGATGAATCA 2588
Db 3478 GGAATCAGAAAGCCCAACAATCTGTATGTGTCAATGCCAAAGCCCAATCCAGCATGAAG 3537
Qy 2589 ACTGGATTACCAAAAGTGTGGGAGCGCTCTTGTGTATCATCGAGGGGGCGGCAACAAC 2648
Db 3538 GCTGAAGAGATAACAACAGGAAGATTCGCACTGCAATGTGTGAGTATATCTGACACAG 3597
Qy 2649 ACCCTGCACCTGCCCTGCTTCGCAAGCATCCGACGCCACATATCTCTCGGTGC 2708
Db 3598 ATAAAGGCAAGGAATATGTTAAGCCCTGTATCATGATGTAGTTGGTCTTGGCAGCTGGCA 3657
Qy 2709 GGCTCCGGTCCCTGGATCACACCGAGGTGCTCGGTGCACTACCCGTATAGGCTTTGGCAT 2768
Db 3658 AAAGACTGGACATCGAGAAATTTGAAGATGGCTCGGTGCAACCTGCAAAATTCATGGCGGAG 3717
Qy 2769 TATCCTTTGACC-----ATCAACTACACATATTTAAATCAGGATGTA 2812
Db 3718 TCTCTTATTGCGGGAGTTTATCTGGGAATCCTGTGAATCGTGACTATATCAGACAGAGA 3777
Qy 2813 CGTGGGAGGGTCAACACACAGGCTGGAAGCTGCTGCAACTGACGCGGGCGCAAGCTTG 2872
Db 3778 CAAGGTGCATTCAGGGATGGAGCCAAAGGAATTTCAAGCCCTCAGGCAACATTCAAAG 3837
Qy 2873 CGATCTGGAGATPAGGACAGGTCCGAGATCGATATGGAGAACATCACATCAGGATTCCT 2932
Db 3838 GATGCTGGATGTACATAGTTGAACATATTTAGTTCACCATCATCAATATGCGGTGTTTGTCT 3897
Qy 2933 AGGACCCCTGCTCGTGTACAGGGGGGTTTTCTTGTGTGAAGA----- 2978
Db 3898 GGGGCCCCCTGATAGGTGTCCACCAACATGCTGTTGTCGGAGGGATGGCTGAATTAGGT 3957
Qy 2979 ---ATCCTCAATACCGAGAGCTAGACTCGTGTGGAGTCTCTCAATTTTCTAGGG 3035
Db 3958 GCCTTCTTTTATCTTCAGGATATGAGGAACAATCATGCTTCAAAAACCTGTGGGC 4017
Qy 3036 GGATCTCCCGTGTGTCTTGCGCCAAAATTCGAGTCCCACTCCCAATCACTCAC----- 3090
Db 4018 ACAGCTGATGAAAAGCTTCGAAAGAAATCATCTTCTATCAATACATACCTCAGACGACA 4077
Qy 3091 CAACCTCTGCTCTCCAAATTTGCTGTGTTATCGCTGGATGTGTCTGGCGGTTTTATCA 3150
Db 4078 CAATCAATGGGAATACAACTGGACCAAGAGATAATTTGTATGTTTATGTTGCTGGGGA 4137
Qy 3151 TATTCTCTTCATCTGCTGCTATGCTCATCTTCTTATTGTTCTTCTGATATCAAG 3210
Db 4138 AAGGAGGAGTGAGAACATTTTCACTCTGGTGATGACATGGATCCAGAGCTTCGTAGCCTG 4197
Qy 3211 GTATGTTGCGCGTTGTCTCTAAATTCAGGATCAACAACAACAGTACGGGACCATGCA 3270
Db 4198 GCTCAGATCTTGATGACCAAGAGTGAAGAAATCTCAACACAGGAACCTATGAATTA 4257
Qy 3271 AAACCTGCAGACTCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAAC 3330
Db 4258 TAAGTACATAAATATATATCAATCACTAATAAGTTAAGAAATACTAATCAATTAGTTA 4317
Qy 3331 CTACGGATGGAATTCGACCTGATTTCCCATCCATCCATCGTCTGGGCTTTCCGAAATACC 3390
Db 4318 ATAAGAAATATAGATTTATTGAATAATCATATAATTAATTAAGTAAAGTAACTAGTATTT 4377
Qy 3391 TATGGGATGGGCTCAGTCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTTCAGT 3450
Db 4378 AGTTAAGTATAGCTAAATTTATATATATGATTTGTACAAATTAATGTAATCATAGACAAAT 4437
Qy 3451 GGTTCGTAGGGCTTTCCCCACCTGTTTGGCTTTTTCAGCTATATGATGA---TGTGGTAT 3507
Db 4438 CACTGCCATGTAATAATCACGGGTATACGGGTGTTTTTATATATGGGGAACAGGGTGGCTT 4497

Qy 3508 GGGGCAAGTCTGTACAGCATCGTGAGTCCCTTTTATACCGCTGT----- 3552
Db 4498 AGGGCAGGTCACTTAAAGTGACCTTTTGTGTATATATGATGTAGATTTCAATTGATC 4557
Qy 3553 -----TACCAATTTTCTTTTGTCTCTGGGTATACATTTAAGAAATTCAGACTCGAGC 3603
Db 4558 GAATACTAATCCTACTGTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4617
Qy 3604 AAGTCTAGAA-----AGCGCGCAAGATATACAGGATCCACTAGCCGTTAGAGC 3653
Db 4618 TGTGCGCAATTTGGCGCGGATCTCGCAATCCCTAGGAGGATTAGGCAAGGGCTTGAGC 4677
Qy 3654 TGGC----- 3657
Db 4678 TCACGCTCTTGTGAGGACAGAAATACAAATCAGGGGAGTATATGAATCTCCATGGAGA 4737
Qy 3658 -----TGATCAGCTTCGACTGTGCTCTTAGTTGCGAGCATCTGTGTGT 3701
Db 4738 AACCCAGATCTACGTATGATCAGCTTCGACTGTGCTCTTAGTTGCGAGCATCTGTGTGT 4797
Qy 3702 TTGCCCCCTCCCGCTGCTTCTTGACCTGGAAGTGCACCTCCACACTGTCTTTCCTA 3761
Db 4798 TTGCCCCCTCCCGCTGCTTCTTGACCTGGAAGTGCACCTCCACACTGTCTTTCCTA 4857
Qy 3762 ATAAATGAGGAATTTGCATCGCATTTGTCTGAGTAGTGTCTATTCTATTCTGGGGGTGG 3821
Db 4858 ATAAATGAGGAATTTGCATCGCATTTGTCTGAGTAGTGTCTATTCTATTCTGGGGGTGG 4917
Qy 3822 GGTGGGCGAGGACAGCAAGGGGAGATTGGGAAGACAATAGCAGCATGCTGGGAGCT 3881
Db 4918 GGTGGGCGAGGACAGCAAGGGGAGATTGGGAAGACAATAGCAGCATGCTGGGAGTGC 4977
Qy 3882 CTT----- 3884
Db 4978 GGTGGGCTCTATGGCTTTTGAGCGGAAAGAACAGCTGGGGCTCGACAGCTCGACTCTA 5037
Qy 3885 ---CCGCTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3941
Db 5038 GAAATGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5097
Qy 3942 CAGCTCACTCAAAGGCGGTAATACGGTTATCCAAGAAATCAGGGGATTAACGAGGAAGA 4001
Db 5098 CAGCTCACTCAAAGGCGGTAATACGGTTATCCAAGAAATCAGGGGATTAACGAGGAAGA 5157
Qy 4002 ACATGTGACAAAAGGCGGACCAAGGCGGAGCAAGCAAGGCGGCGGCTGCTGGCGT 4061
Db 5158 ACATGTGACAAAAGGCGGACCAAGGCGGAGCAAGGCGGAGCAAGGCGGCGGCTGCTGGCGT 5217
Qy 4062 TTTTTCATAGGCTCGCGCCCTCGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGGT 4121
Db 5218 TTTTTCATAGGCTCGCGCCCTCGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGGT 5277
Qy 4122 GCGAAACCCGACAGACTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTGC 4181
Db 5278 GCGAAACCCGACAGACTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGTGC 5337
Qy 4182 GCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAA 4241
Db 5338 GCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAA 5397
Qy 4242 GGTGCGGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCT 4301
Db 5398 GGTGCGGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCGCT 5457
Qy 4302 CCAAGCTGGGCTGTGTGACAGAACCCCCGTTTTCAGCCGACCGCTCGGCTTATTCGGTA 4361
Db 5458 CCAAGCTGGGCTGTGTGACAGAACCCCCGTTTTCAGCCGACCGCTCGGCTTATTCGGTA 5517
Qy 4362 ACTATGCTTGTAGTCCAAACCGGTTAAGACAGACTTATCGCCACTGGGAGGAGCCACTG 4421
Db 5518 ACTATGCTTGTAGTCCAAACCGGTTAAGACAGACTTATCGCCACTGGGAGGAGCCACTG 5577

```
QY 4422 GTAACAGGATTACAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGC 4481
DB |||||||
QY 5578 GTAACAGGATTACAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGC 5637
DB |||||||
QY 4482 CTAACCTACCGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTA 4541
DB |||||||
QY 5638 CTAACCTACCGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTTTA 5697
DB |||||||
QY 4542 CCTTCGGAAGAGAGTTGGTAGCTCTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTG 4601
DB |||||||
QY 5698 CCTTCGGAAGAGAGTTGGTAGCTCTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTG 5757
DB |||||||
QY 4602 GTTTTTTTGTTTCAGAGCAGAGATTACCGGCGAGAAAAAGAGATCTCAAGAGATCCTT 4661
DB |||||||
QY 5758 GTTTTTTTGTTTCAGAGCAGAGATTACCGGCGAGAAAAAGAGATCTCAAGAGATCCTT 5817
DB |||||||
QY 4662 TGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGG 4721
DB |||||||
QY 5818 TGNATCTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGG 5877
DB |||||||
QY 4722 TCATGAGATTATCAAAAGAGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAA 4781
DB |||||||
QY 5878 TCATGAGATTATCAAAAGAGATCTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTAA 5937
DB |||||||
QY 4782 AATCAATCTAAGTATATAGTAAACTTGGTCTGACAGATTACCAATGCTTAATCAGTG 4841
DB |||||||
QY 5938 AATCAATCTAAGTATATAGTAAACTTGGTCTGACAGATTACCAATGCTTAATCAGTG 5997
DB |||||||
QY 4842 AGGCACCTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTC 4894
DB |||||||
QY 5998 AGGCACCTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTC 6050
DB |||||||
```

RESULT 5

```
US-10-491-121-24
; Sequence 24, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 6148
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-VP35
US-10-491-121-24
```

```
Query Match 40.7%; Score 2395.6; DB 18; Length 6148;
Best Local Similarity 71.6%; Pred. No. 0;
Matches 3516; Conservative 0; Mismatches 1194; Indels 204; Gaps 20;
```

```
QY 1 TCGCGGTTTCGGTGATCGGTGAAACCTCTGACACATGACGTCCCGGAGACGGTCA 60
DB |||||||
QY 61 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGGGTTCAGCGGGTG 120
DB |||||||
QY 61 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGCGGGTTCAGCGGGTG 120
DB |||||||
```

```
QY 121 TTGGCGGCTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB |||||||
QY 121 TTGGCGGCTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
DB |||||||
QY 181 ACCATATGAAGCTTTTTTGCAAAAGCCTAGGCCTCCAAAAAGCCTCCTCACTACTTCTGG 240
DB |||||||
QY 181 ACCATATGCGG-----TGTGAATACCGCACA 207
DB |||||||
QY 241 AATAGCTCAGAGCCGAGCGCCCTCGGCCTCTGCATAAATAAAAAAATAGTCAGCCA 300
DB |||||||
QY 208 GATGCGTAAGGAGAAAAATACCGCATCAG----- 235
DB |||||||
QY 301 TGGGCGGAGAAATGGCGGAACTGGCGGGGAGGAGAAATATTGGCTATTGGCCATTGCAT 360
DB |||||||
QY 236 -----ATTGGCTATTGGCCATTGCAT 256
DB |||||||
QY 361 AGCTTGATCTATATCATATAATGTACATTTATATTGGCTCATGTCCAATATGACCGCCA 420
DB |||||||
QY 257 AGCTTGATCTCATATCATATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCA 316
DB |||||||
QY 421 TGTGACATTGATTATTGACTAGTTATTAATAGTAAATCAATTAACGGGGTCAATTAGTTTCA 480
DB |||||||
QY 317 TGTGACATTGATTATTGACTAGTTATTAATAGTAAATCAATTAACGGGGTCAATTAGTTTCA 376
DB |||||||
QY 481 AGCCCATATATGAGGTTCCGCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCG 540
DB |||||||
QY 377 AGCCCATATATGAGGTTCCGCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCG 436
DB |||||||
QY 541 CCCAAGCAGCCCGGCCATTGACGTCATTAATAGACGTATGTTCCCATAGTAGCCCAATA 600
DB |||||||
QY 437 CCCAAGCAGCCCGGCCATTGACGTCATTAATAGACGTATGTTCCCATAGTAGCCCAATA 496
DB |||||||
QY 601 GGGACTTTCCATTGACGTCATTAATGGGTGGAGTATTTACGGTAAACTGCTCCCACTGGCAGTA 660
DB |||||||
QY 497 GGGACTTTCCATTGACGTCATTAATGGGTGGAGTATTTACGGTAAACTGCTCCCACTGGCAGTA 556
DB |||||||
QY 661 CATCAAGTGTATCATATGCAAGTCCGCGCTTATGAGTCATCAATGACGTAATGCGCC 720
DB |||||||
QY 557 CATCAAGTGTATCATATGCAAGTCCGCGCTTATGAGTCATCAATGACGTAATGCGCC 616
DB |||||||
QY 721 GCCTGGCATTTATGCCAGTACATGACCTTTACGGGACTTTCTTACTTTGGCAGTACATCTAC 780
DB |||||||
QY 617 GCCTGGCATTTATGCCAGTACATGACCTTTACGGGACTTTCTTACTTTGGCAGTACATCTAC 676
DB |||||||
QY 781 GTATTAGTCATCGCTATTACATGATGCGGTTTTTGGCAGTACACCAATGGGCGTGA 840
DB |||||||
QY 677 GTATTAGTCATCGCTATTACATGATGCGGTTTTTGGCAGTACATCAATGGGCGTGA 736
DB |||||||
QY 841 TAGCGGTTTGTACTCAGCGGATTTTCCAGTCTCCACCCATTGACGTCATGAGGAGTTG 900
DB |||||||
QY 737 TAGCGGTTTGTACTCAGCGGATTTTCCAGTCTCCACCCATTGACGTCATGAGGAGTTG 796
DB |||||||
QY 901 TTTTGGCACCAAAATCAACGGGACTTTCCAAATGTCTGTAATAACCCCGCCCGTTGACG 960
DB |||||||
QY 797 TTTTGGCACCAAAATCAACGGGACTTTCCAAATGTCTGTAATAACCCCGCCCGTTGACG 856
DB |||||||
QY 961 CAAATGGCGGTAGCGGTGTACGTTGAGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC 1020
DB |||||||
QY 857 CAAATGGCGGTAGCGGTGTACGTTGAGGAGTCTATATAAGCAGAGCTCGTTTAGTGAAC 916
DB |||||||
QY 1021 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGAC 1080
DB |||||||
QY 917 CGTCAGATCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGAC 976
DB |||||||
QY 1081 CGATCCAGCTCCGCGGCGGAAACGGTGCATTGGAAACCGGATTCCTCCGTCGCAAGAGT 1140
DB |||||||
QY 977 CGATCCAGCTCCGCGGCGGAAACGGTGCATTGGAAACCGGATTCCTCCGTCGCAAGAGT 1036
DB |||||||
QY 1141 GACGTAAGTACCGCTATAGACTCTATAGGACACACCCCTTTGGCTCTTATGCTATGCTATA 1200
DB |||||||
QY 1037 GACGTAAGTACCGCTATAGACTCTATAGGACACACCCCTTTGGCTCTTATGCTATGCTATA 1096
DB |||||||
QY 1201 CTGTTTTTGGCTTGGGGCTATACACCCCGC-TCCTTATGCTATAGGTGATGGTATAGC 1259
DB |||||||
```

1097 Db |||||CTGTTTGGCTTGGGCTTATACACCCCGCTTCTTATGCTATAGGTGATGCTATAGC 1156
1260 QY TTAGCCTATAGGTGGTGTATTTGACCATTTATGACCACTCCCTTATTTGGTGACGATACT 1319
1157 Db TTAGCCTATAGGTGGTGTATTTGACCATTTATGACCACTCCCTTATTTGGTGACGATACT 1216
1320 QY TTCCATTACTTAATCCATAACATGGCTCTTTGGCCACAACATCTCTATTTGGCTATATGCCA 1379
1217 Db TTCCATTACTTAATCCATAACATGGCTCTTTGGCCACAACATCTCTATTTGGCTATATGCCA 1276
1380 QY ATACTCTGTCTTACAGAGCTGACAGGACTCTGTATTTTACAGAGATGGGGT-CAATTT 1438
1277 Db ATACTCTGTCTTACAGAGCTGACAGGACTCTGTATTTTACAGAGATGGGGTCCCAATTT 1336
1439 QY ATTATTTTACAAATTCACATATACAAACCGCGCTCCCGGTGCCCGAGTTTTTATTA 1498
1337 Db ATTATTTTACAAATTCACATATACAAACCGCGCTCCCGGTGCCCGAGTTTTTATTA 1396
1499 QY CATAGCGTGGGATCTCC--GACATCTCGGGTACGTTTCCGGACATGGGCTCTTCTCCG 1555
1397 Db CATAGCGTGGGATCTCCACGCGAATCTCGGGTACGTTTCCGGACATGGGCTCTTCTCCG 1456
1556 QY GTAGCGCGGAGCTTCCACATCCGAGCCTGGTCCCATCGTCCAGCGGCTCATGTGCG 1615
1457 Db GTAGCGCGGAGCTTCCACATCCGAGCCTGGTCCCATCGTCCAGCGGCTCATGTGCG 1516
1616 QY TCGGAGCTCTTGTCTTAAACAGTGGAGGCGAGACTTGGGACACACAATGCCACCA 1675
1517 Db TCGGAGCTCTTGTCTTAAACAGTGGAGGCGAGACTTGGGACACACAATGCCACCA 1576
1676 QY CCACAGTGTGCCGACAAAGCGCGTGGGTAGGATATGTCTGAAATAGCTCGGAG 1735
1577 Db CCACAGTGTGCCGACAAAGCGCGTGGGTAGGATATGTCTGAAATAGCTCGGAG 1636
1736 QY ATTGGGCTGCGAC-CTGGAGCGAGATGGAAGACTTAAGGAGCGGCGAGGAAGATGCGAG 1794
1637 Db ATTGGGCTGCGACGCGGTGACGAGATGGAAGACTTAAGGAGCGGCGAGGAAGATGCGAG 1696
1795 QY GCAGCTGAGTTGTGTATCTGATAAGAGTACAGAGTAACTCCCGTTGGGTGCTGTAA 1854
1697 Db GCAGCTGAGTTGTGTATCTGATAAGAGTACAGAGTAACTCCCGTTGGGTGCTGTAA 1756
1855 QY CGGTGAGGCGAGTGTAGTCTGACGAGTACTCGTTCGCGCGCGCCACACAGACATA 1914
1757 Db CGGTGAGGCGAGTGTAGTCTGACGAGTACTCGTTCGCGCGCGCCACACAGACATA 1816
1915 QY ATAGCTGACAGACTAACAGACTGTTCTTTCATAGGGTCTTTTCTGAGTCAACGTCGTC 1974
1817 Db ATAGCTGACAGACTAACAGACTGTTCTTTCATAGGGTCTTTTCTGAGTCAACGTCGTA 1876
1975 QY GAC---GAATTCAGCAATCATGGATGCAATGAAGAGGGGCTCTGCTGTGCTGCTG 2030
1877 Db TTCTCTAGACCTCGAAGCTTATGTCTTCAATGAAAGAAAGCTGGTCTAACAGATG 1936
2031 QY CTGTGTGGAGCAGTCTTCTGTTTCGCGCAGCGCTAGCGAAACCCACGTC-ACCGGGGGAAG 2089
1937 Db ACAACTAGAAACAAAGGCGAGGSCCATCTGCGGCCACGACTCAAAACGACAGATGCCA 1996
2090 QY TGGCGGCCACAGTGTCTGTGATTTGTATGCTTCTGCAAGGGTCTTTTCTGAGTCAACGTCG 2149
1997 Db GGCCTTGAGCTTTTCGGGCTGGATCTCTGAGCAGCTAATGACCGGAAGAAATTTCTGTAAAGC 2056
2150 QY CCAGCTGATCAACACCAACGAGTTGGGACCTCAATAGCAGGCGCTGACTGGAATGCAATGA 2209
2057 Db GACATCTTCTGTGATATTGAGAAACAATCCAGGATTTATGCTAGCAGTCCCAATGCAACAA 2116
2210 QY TAGCCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTCTCAACTCTTCAGG 2269
2117 Db ACGAAGCCAAACCGAAGACGCGCA---ACAGTCAAAACCAACGACCAATTTGCAAT 2173
2270 QY CTGTCTGAGAGGCTAGCCAGCTGCGGACCCCTTACCGATTTTTCAGCAGGGCTGGGGCCC 2329

2174 Db CATAGTTTTTGAGG-----AGGTAGTACAAACATATGGCTTCTTATTTGGCTACTGTTGTGCAAC 2228
2330 QY TATCAGTTATGCAACGGAAGCGGCCCCGACAGCGCCCCCTACTGCTGGCACTACCCCC 2389
2229 Db AACAAACCATCGCATCAGAAATCATTTAGAACACGCAATTAGAGTCTTGAGAAATGGTCTAA 2288
2390 QY AAAACCTTCCGGTATTTGTCCCGCGAAGAGTGTGTGGTCCGGTATATTTGCTTCACTCC 2449
2289 Db AGCCAGTTTATGATATGCAAAACAACTCTCTCATTTGAACAGGGTTTGTGCTGAGATGG 2348
2450 QY CAGCCCCGTGGTGGGGAACGACCGACAGGTGCGGCGCGCCACCTACAGCTGGGGTGA 2509
2349 Db TTGCAAAATATGATCTTTCTGGTGAACAAACCGGTGCGGCAACCGCTGCGGCAA 2408
2510 QY AAATGATACGAGCTTCTGCTTAAATAACAGGCGCACCGCTGGGCAATTTGGTTGCG 2569
2409 Db CTGAGG-----CTTATTTGGCGGACATGCTCAACACCTCGGACCATCATCTTTAT 2461
2570 QY TTGTACTCGATGAACCTCAACTGGATTACCAAAGTGTGGGAGCGCTCTCTTGTGTAT 2629
2462 Db GAAGAAAGTGCATTCGCGGTAAAGATTGAATCTAGAGATGAGACCGTCC---TCAAAG 2517
2630 QY CCGAGGGGCGGCAACACACCTGCACTGCCCACTGATTTGCTTCGCAAGCATCGGA 2689
2518 Db TGTTAGGGAGGCAATTCACCAATCTAAACAGTACCCT----- 2554
2690 QY CGCCACATACTCTCGTGGCGCTCCGCTCCCTGGATCACACCGAGTGCCTGCTCGACTA 2749
2555 Db --TCATACTAGGAAATTTTGGGAAACCTTGACATTTTCGGCAAGGATTTGAGAAACA 2612
2750 QY CCCGTATAGGCTTTGGCATTTATCTTGTATACCATCACTCAACCATATTTAAATTCAGGAT 2809
2613 Db TTATGTATGATCATTGCTGCTGTTTGGAACTGCTTTCCACCAATTTAGTACAGTATTT 2672
2810 QY GTAAGTGGAGGGTGGAAACAGGCTGGAAGTGCCTCAACTGCACTGCAAGCGGGGCGAAG 2869
2673 Db GTAAATTTGGGAAAGATAGCAACTCATTTGGACATCATTTGCTGATTTCCAGGCCAGCC 2732
2870 QY TTGGCATCTGGAAGATAGGACAGGTCGAGATCGATATGGAGAACATCACATCAGGAT 2929
2733 Db TGGCTGAAGGAGACTCTCTCTCAATGTGCTCAATTTCAAAATTACAAAAGAGTTCCAATCT 2792
2930 QY CTTAGGACCCCTGCTCGTGTATTACAGCGGGGTTTTTCTTTGTGACAAGAAATCTCTCAAT 2989
2793 Db TCCAAAGATGCTCTCATCTGTCATCCACATCCGCTTTTCAGGTGACATTTCCCGAGCTT 2852
2990 QY ACCGAGAGTCTAGACTCGTGTGGA-CTTCTCTCAATTTTTCAGGGGATCTCCCGTGT 3048
2853 Db GCCAGAAAGCTTGGCTCCAGTCCACCATCCCAAGATTGATCGAGGTGGGATGTGT 2912
3049 QY GTCTTTGGCCAAATTCG---CAGTCCCAACCTCCAATCACTCACCAACCTCTGCTCCTC 3105
2913 Db TTTTCAGCTTCAAGATGTTAAACACTTGGACTCAAAATTTGAGCCAAATCTCTCTTCCCT 2972
3106 QY CAATTTGCTCTGTTATCGCTGGATGTGCTCGCGGTTTTTATCATATTCTCTTTCATCC 3165
2973 Db CGAAAGAGGCGCAATAATAGCAGAGGCTTCAACTGCTGAACATATAGGTACGTTACAT-- 3030
3166 QY TCGTGTATGCCCTCATCTTCTTATTTGGTTCCTGATTTATCAAGTATTTGTCGCCGTTT 3225
3031 Db -----TAATGATACACTTGTGAGTATCAGCCCTGGATAATAAGTCAATTTAAACGACC 3084
3226 QY GTCTCTAATTCAGGATCAACAAACAGTACGGGACCATGCAAAACCTGCACGACTC 3285
3085 Db AAGATAAAATTTGTTTCAATATCTCGTAGCAGCTTAAATAATAATGTAATAGAGCTATAT 3144
3286 QY CTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGTGTTGTTGTAACAAACCTACGGATGGAAT 3345
3145 Db CTCTGACAGGGGATCCAGATCTGCTGTGCTTCTAGTTGCCAGC----- 3189
3346 QY GCACCTGTATTTCCCATCCATCGTCTGGGCTTTTCGCAAAATACCTATTCGGAGTGGGCT 3405
3190 Db ---CATCTGTTGTTTGGCCCTCCCGCTGCTTCTTGTGACCTGGAAGGTGCCACTCCCA 3246

Qy	181	ACCATATGACGCTTTTTCGAAAAAGCCTAGGCCTCTCAAAAAGCCTCTCTCACTACTTCTCGG	240
Db	181	ACCATATGCGG-----TGTAATAATACCGCAC	207
Qy	241	AATAGCTCAGAGCGGAGCGGCTCGGCTCTGCATAAATAAAAAAATTAGTCAGCCCA	300
Db	208	GATCGGTAAAGGAGAAATACCGCATCAG-----	235
Qy	301	TGGGGCGGAGAAATGGCGGMACTCGGCGGGAGGGAATTAATTGGCTATTGGCCATTGGCAT	360
Db	236	-----ATTGGCTATTGGCCATTGGCAT	256
Qy	361	ACGTTGTATCTATATCATATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCA	420
Db	257	ACGTTGTATCCATATCAATAATGTACATTTATATTGGCTCATGTCCAACATTTACCGCCA	316
Qy	421	TGTTGACATTCATTATTGACTAGTATTATTAATAGTAATCAATTACGCGGTTCATTAGTTTCAT	480
Db	317	TGTTGACATTGATTATTGACTAGTATTATTAATAGTAATCAATTACGCGGTCAATTAGTTTCAT	376
Qy	481	AGCCCATATATGAGTATTCGCGTTTACATAACTTACGGTAAATGGCCCGCCCTGGCTGACCG	540
Db	377	AGCCCATATATGAGTATTCGCGTTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCG	436
Qy	541	CCCAACGACCCCGCCCATTTGACGTCAATAATAGACGTATGTTGCCATAGTAACGCCAATA	600
Db	437	CCCAACGACCCCGCCCATTTGACGTCAATAATAGACGTATGTTGCCATAGTAACGCCAATA	496
Qy	601	GGGACTTTCCATTCACGCTCAATGGTGGAGTATTTACGTTAACTGACGTCGCCACTTGGCAGTA	660
Db	497	GGGACTTTCCATTCACGCTCAATGGTGGAGTATTTACGTTAACTGACGTCGCCACTTGGCAGTA	556
Qy	661	CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCGAATGACGGTAAATGGCCCC	720
Db	557	CATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCGAATGACGGTAAATGGCCCC	616
Qy	721	GCCTGGCATTTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC	780
Db	617	GCCTGGCATTTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC	676
Qy	781	GTATTAGTCATCGCTATTACCATGGTGCATGCGGTTTTGGCAGTACACCAATGGCGTGA	840
Db	677	GTATTAGTCATCGCTATTACCATGGTGCATGCGGTTTTGGCAGTACATCAATGGCGTGA	736
Qy	841	TAGCGGTTTGACTCAGCGGGATTTCCAAGTCTCCACGCCCATTTGACGTCGAATGGAGTTTG	900
Db	737	TAGCGGTTTGACTCAGCGGGATTTCCAAGTCTCCACGCCCATTTGACGTCGAATGGAGTTTG	796
Qy	901	TTTTGGCACCAAAATCAACGGGACTTTTCCAAAATGTGTAATAACCCGCCCGCTTGAACG	960
Db	797	TTTTGGCACCAAAATCAACGGGACTTTTCCAAAATGTGTAATAACCCGCCCGCTTGAACG	856
Qy	961	CAATGGGCGGTAGCGGTGTACGTTGGAGGTCTATATAGCAGAGCTCGTTTAGTGAAC	1020
Db	857	CAATGGGCGGTAGCGGTGTACGTTGGAGGTCTATATAGCAGAGCTCGTTTAGTGAAC	916
Qy	1021	CGTCAGATCGCTCGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGAC	1080
Db	917	CGTCAGATCGCTCGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGAC	976
Qy	1081	CGATCCAGCTTCGCGGCGGGAAACGTTGCATTTGAAACGCGGATTCGCCGTGCCAAGAGT	1140
Db	977	CGATCCAGCTTCGCGGCGGGAAACGTTGCATTTGAAACGCGGATTCGCCGTGCCAAGAGT	1036
Qy	1141	GACGTAACTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGTCATGCTATA	1200
Db	1037	GACGTAACTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGTCATGCTATA	1096
Qy	1201	CTGTTTTTGGCTTGGGCTATACACCCCGC-TCCTTATGCTATAGTGTATGATATAGC	1259
Db	1097	CTGTTTTTGGCTTGGGCTATACACCCCGCTTCTTATGCTATAGTGTATGATATAGC	1156
Qy	1260	TTAGCCTATAGGTGGGTTATTAGACCATTTATGACCACTCCCCCTATTTGGTAGCATPACT	1319

1157 TTAGCCCTATAGGTGTGGGTATTGACCATTATTGACCACCTCCCTATTGGTAGCATACT 1216

1320 TTCCATTACTTAATCCATAACATGGCTCTTTGGCACAACTATCTCTATTGGCTATATGCCA 1379

1217 TTCCATTACTTAATCCATAACATGGCTCTTTGGCACAACTATCTCTATTGGCTATATGCCA 1276

1380 ATACTCTGTCTCTTCAGAGACTGACACGGACTCTGTATTTTTTACAGGATGGGTF - COATTT 1438

1277 ATACTCTGTCTTCAGAGACTGACACGGACTCTGTATTTTTTACAGGATGGGTFCCCATTT 1336

1439 ATTATTTACAAATTCACATATACAAACGCGTCCCCCGTGGCCGCGAGTTTTTATTATAA 1498

1337 ATTATTTACAAATTCACATATACAAACGCGCGTCCCCCGTGGCCGCGAGTTTTTATTATAA 1396

1499 CATAGCGTGGGATCTCC --- GACATCTCGGGTAGGTGTTCCGACATAGGCTCTTCTCCG 1555

1397 CATAGCGTGGGATCTCCACGCGAATCTCGGTACGTGTTCCGACATAGGCTCTTCTCCG 1456

1556 GTAGCGGGAGGCTTCCACATCCGAGCCCTGTFCCCATCCGTCACGCGCTCATGTGTCG 1615

1457 GTAGCGGGAGGCTTCCACATCCGAGCCCTGTFCCCATGCCCTCAGCGCTCATGTGTCG 1516

1616 TCGGAGCTCTTGTCTCTTAAAGTGGAGCCAGACTTAGGCACAGCACAAATGCCACCA 1675

1517 TCGGAGCTCTTGTGCTCTTAAAGTGGAGCCAGACTTAGGCACAGCACAAATGCCACCA 1576

1676 CCACAGTGTGGCGCAAGGCGTGGCGGTAGGGTATGTGTCGAAATGAGCTCGGAG 1735

1577 CCACAGTGTGGCGCAAGGCGTGGCGGTAGGGTATGTGTCGAAATGAGCTGGAG 1636

1736 ATTGGGCTCGAC - CTGGACGCAGATGGAAGACTTAAAGCAGCGGAGAGAGATGCGAG 1794

1637 ATTGGGCTCGACCGGCTGACGCAGATGGAAGACTTAAAGCAGCGGAGAGAGATGCGAG 1696

1795 GCAGCTGAGTTGTTGTTCTGATAAGAGTCAAGGTAACTCCCGTGGCGGTGCTGTTAA 1854

1697 GCAGCTGAGTTGTTGTTCTGATAAGAGTCAAGGTAACTCCCGTGGCGGTGCTGTTAA 1756

1855 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCTGTCGCCGCGCGCCACACAGACATA 1914

1757 CGGTGAGGGCAGTGTAGTCTGAGCAGTACTCTGTCGCCGCGCGCCACACAGACATA 1816

1915 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGCTTTTTTGCAGTCAACCGTCGTC 1974

1817 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGCTTTTTTGCAGTCAACCGTCGTC 1876

1975 GAGCAATTCAAGCAATCATGGATGC --- AATGAAGAGAGGGCTGTGCTGTGCTGCTG 2030

1877 GACACGTGTGATCAGATATCGGGCGGCTCTAGAATTTAGGATTTGGGCTTTTAGAGATTTC 1936

2031 CTGTGTGGAGCAGTCTTCGTTTCCCGACGCGTAGCGAAACCAAGTCAACCGGGGAGT 2090

1937 ACTACTAGTTAGGAATTCCTAAATCATGGGGCAGATTATTAATCTTTTCAAGAGTGCC 1996

2091 GCGGCCACACTGTGTCTGGATTTGTTAGC - CTCCTCGCACAGGGCGCCAGCAGAACGT 2149

1997 ACATGTAATAGAGGAAGTCAAGAACATGTGCTAATTGGCTTTTCTCTATTGGCAATCTT 2056

2150 CCAGCTGATCAACCAACGGCAGTTGGCACTTCAATAGCAAGGCCCTGAACTGCAATGA 2209

2057 GAAGGGCTTGTATAACATCGCTACATGTGGGATTTATGGAATTTGGTTGGCTTTTTTATCTT 2116

2210 TAGCCTCAACACGGCTGTTGGCAGGGCTTTTCTATCACCAACAAGTTCACCTCTTCAGG 2269

2117 GTGTGCAAGTCTTGTTCCTTAACCCCTTAAGGGGGATATGAGTCTCAAACTTTAGATT 2176

2270 CTGTCTCGAGAGGCTAGCCAGCTGCGCAGCCCTTACCGATTTTTCAGAGGGCTGGGGCCC 2329

2177 AATATGAGAGCCCTTAAACATGACCATGCCCTTATCATGCACCAAGAAACAGCAGTATCA 2236

2330 T --- ATCAGTTATGCCAAACGGAAGCGGCCCGGACACAGCGC --- CCCTACTGTGCGCACTA 2383

Db 2237 TTACATAAGAGTGGGCAATGAGACTGGATTAGATTGACTTTAACTAAACACAGCAATTAT 2296
Qy 2384 CCCCCAAAACCTTGGCGTATTGTGCCCGCGAAGAGTGTGTGGTCCGGTATATTGCTT 2443
Db 2297 AAATCACAATTTTGGCAACTTATCCGATGCTCACAAAAAGAAATCTTTATGATCATGCTCT 2356
Qy 2444 CACTCCAGCCCGTGGTGGGAGACGACGACAGGTGGGGCGGCCACCTACAGCTG 2503
Db 2357 CATGAGCATCATCTCAACATTCATCTATCCATTCCTCAATTCGATGATGAAGCCAT 2416
Qy 2504 GGGTGAATGATACGGAGCTCTTCGTCTTAACAATACAGGCCACCGCTGGGCAATTG 2563
Db 2417 GAGTTGTGATTTCAATGGAGGAATCAGTGTGCATACAACTCTCATCTCTATGCG 2476
Qy 2564 GTTCGGTTGTACATGGATGAATCAACTGGATTACCAAAAGTGTGGAGCGCTCTCTTG 2623
Db 2477 TGGGATCGGCGGAACACTGTGGGACAGTTCGCAACGGAGTGTTCGCAAACTTTATGAG 2536
Qy 2624 TGTATCGAGGGGGGCAACAAACCTGCACTGCCCCCACTGATTGCTTCGCAAGCA 2683
Db 2537 AATGGCTTGGGTGAAGATACATTGCAATTAGACTCAGGAAAGGGAACCTGGGACTGTAT 2596
Qy 2684 TCCGAGCGCCACATACTCTCGTGGGCTCCGTCCTCGGATCACACCCAGGTGCTGCT 2743
Db 2597 AATGACCACTACCACTGCTGATTAATCAAAATACAACTGGGAGGACCTGCCAATT 2656
Qy 2744 CGACTACCGGTATAGCTTTTG--CATTAATCTTTGATCAATCAACTACACCATATTTAAA 2801
Db 2657 CTCAGACGCTCTCTATCCGCTGACCTTGGCTTTTGTCAAAAGGACAAAGAGATATA 2716
Qy 2802 ATCAGATGTAGTGGAGGGTGCNAACACAGCTGGAAGCTGCCTGCAACTGGAGCGG 2861
Db 2717 TATAAGTAGGAGCTCTTGGGACCTTCAC-----CTGGCAATTTGTGAGATTCTG 2766
Qy 2862 GCGAAGCTTGCATCTGAAGATAGGACAGTCCGAGATCGATATGAGAAATACACA 2921
Db 2767 AGGCAATGAACACCACTGCTGTTATTTAAACAGGTGGATGCTAATTGAACGAGAC 2826
Qy 2922 TCAGGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTTCTTGTGTAACAAGAATC 2981
Db 2827 TCAAGTGTTTTGGGAATACAGCTGTGGCAAAATGCAATGAGAAGCATGATGAGGAGTTT 2886
Qy 2982 CTCACATACCGAGAGTCTAGACTCGTGGTGAATCTCTCAATTTTCTAGGGGATCT 3041
Db 2887 GTGACATGTGAGATTGTTGATTTCAACAAGCAAGCAATCCGTAGTTGAAGGCTGAGG 2946
Qy 3042 CCGGTGTCTTGGCCAAAATTCGAGTCCCAACCTCCAACTCACTCAACCACTCTCTGT 3101
Db 2947 CCCAGATGAGTATCAATTAATAATAAGCCGTGNAATGCCCTTAATCAATGATCAATTAA 3006
Qy 3102 CTCCA-----ATTGTCTCGTTATCGCTGGATGTGTCTCGGCGTTTAT 3148
Db 3007 TCATGAAGAACCAATTAAGAGACATCATGGCAATTCCTACTGCAATTAACAGCAAGTAT 3066
Qy 3149 CATATTCTCTTCATCTGCTGCTATGCTCATCTCTTATTGGTTCTTCTGGAATTACA 3208
Db 3067 GGTACCTTAATCATACTAGTAGGGGGAACATCACTACAAAGTGTGTGGCTTATATCCA 3126
Qy 3209 AGGTATGTTGGCGTTGTCTTAATTCAGGATCAACAACAACAGTACGGGACCAATG 3268
Db 3127 ATGGGTCAATCTAATTAAGAACCCAGTCTCTGTATGACATAGNAACGAAAGCCGACAATA 3186
Qy 3269 CAAAACCTGCAAGACTCTG-----CTCAAGGCAACTCTATGTTTCCTCATGTT 3318
Db 3187 TGATCACAGAGATGCTTCAGAAAGAAATACATTGAAGACAAAGGAAAAAGCCCTTGTAGG 3246
Qy 3319 GCTGTACAAAACCTACGGATGGAATTGC---ACCTGATTTCCCATCCCATCGTCTGGG 3375
Db 3247 GATCCAGATCTGCTGTGCTCTTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCCGTGC 3306
Qy 3376 CTTTCGAAAAATACCTATATGGAGTGGGCGCTCAGTCCGTTTCT-CTTGGCTCAGTTTACTA 3434
Db 3307 CTTCTTGACCTGGNAGTGGCACTCCCACTGCTCTTCTTAATAAATAGGAAATTG 3366

Qy 3435 GTGCCAATTTGTTCACTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCACTATATGG 3494
Db 3367 CATCGCAATTTCTGAGTAGGTCTATTCTATTCTGGGGGTGGGTGGGCGAGGACAGCA 3426
Qy 3495 ATGATGTGGTATTTGGGGCCCAAGTCTGTACAGCATCGGTGAGTCCCTTTATATCCGCTGTTA 3554
Db 3427 AGGGGAGGATTTGGGAAGACAATAGCAGCATGCTGGGGATGCGGTGGGCTCTATGGGTA 3486
Qy 3555 CCAATTTTCTTTTG---TCTCTGGGTATACATTTAAGAAATTCAGACTCGAGCAAGTCTAG 3611
Db 3487 CCCAGGTCTGAAGAAATGACCCGCTTCTCTCGGCCAGAAAGACGAGGACACATCCCC 3546
Qy 3612 AAAGCGGCCCAAGATATCAAGGATCCACTAGCGCTTTAGAGCTCGCTGATACGCTCGAC 3671
Db 3547 TTCTCTGTGACACACCTGTCCACGCCCTGGTCTTTAGTTCCAGCCCCCACTCATAGGAC 3606
Qy 3672 -----TGTGCTTCTAGTTGCGAGCAATCTGTTGTTTTCGCCCTCCCCCGTGC 3719
Db 3607 ACTCATAGTCTAGGAGGGTCCGCCCTTCAATCCCAACCCGCTAAAGTACTTTGGAGCGTCT 3666
Qy 3720 TTCTCTTGAACCTTGAAGGTGCCACTCCCACTGCTCCTTTCTTAATAAAATGAGGAAATTCG 3779
Db 3667 CTCCCTCCCTCATCAGCCCAACCAACCACTAGCTCCAGAGTGGGAAGAAATTTAA 3726
Qy 3780 ATCGCATTTGTCTGAGTAGTGTCTATTCTTCTGGGGGTGGGGTGGGCGACGACAGCA 3839
Db 3727 GCAAGATAGGCTATTAAAGTGCAGAGGAGAGAAAATGCTTCCAACTGTGAGGAGATAT 3786
Qy 3840 GGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGAGCTCTTCGGCTTCTCGCTCA 3899
Db 3787 GAGAAATCATAGAAATTTTAAAGCCCATG--GCCTTAATCTTCGGCTTCTCGCTCA 3844
Qy 3900 CTGACTCGCTCGCTCGTTCGGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGG 3959
Db 3845 CTGACTCGCTCGCTCGTTCGGTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGCGG 3904
Qy 3960 TAATACGGTTATTCACAGAAATCAGGGATTAACGAGGAAAGAAACATGTGAGCAAAAGCC 4019
Db 3905 TAATACGGTTATTCACAGAAATCAGGGATTAACGAGGAAAGAAACATGTGAGCAAAAGCC 3964
Qy 4020 AGCAAAAGCCAGGAAACCGTAAAAAGCGGCTTCGCTGGCGTTTTTTCATAGGCTCCGCG 4079
Db 3965 AGCAAAAGCCAGGAAACCGTAAAAAGCGGCTTCGCTGGCGTTTTTTCATAGGCTCCGCG 4024
Qy 4080 CCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTCAAGTGGCGAAACCCGACAGGAC 4139
Db 4025 CCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGTGGCGGAAACCCGACAGGAC 4084
Qy 4140 TATAAAGATACAGGCGTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCC 4199
Db 4085 TATAAAGATACAGGCGTTTCCCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCC 4144
Qy 4200 TGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAAT 4259
Db 4145 TGCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT 4204
Qy 4260 GCTACGCTGTAGGTATCTCAAGTTCGGTGTAGTTCGGTTCGCTTCAAGTGGCGCTGTGTC 4319
Db 4205 GCTACGCTGTAGGTATCTCAAGTTCGGTGTAGTTCGGTTCGCTTCAAGTGGCGCTGTGTC 4264
Qy 4320 AGAAACCCCGCTTCAGCCCGACCGCTGGCTTATCCGGTAACTATCGTCTTGAGTCCA 4379
Db 4265 AGAAACCCCGCTTCAGCCCGACCGCTGGCTTATCCGGTAACTATCGTCTTGAGTCCA 4324
Qy 4380 ACCCGTGAAGACACGACTTATCGCACTGGGAGAGCCACTGTAACAGGATTAGCAGAG 4439
Db 4325 ACCCGTGAAGACACGACTTATCGCACTGGGAGAGCCACTGTAACAGGATTAGCAGAG 4384
Qy 4440 CGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGGCTTAACACGCTACACTA 4499
Db 4385 CGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGTGGGCTTAACACGCTACACTA 4444

3081	AGTTTGTA	CAAAAAAGCAGGCTCCGCGCCGCAACCATGGCAAGTGTGTGTGGTTGGGGGCC	3140	
Db				
1176	CCCTTTGGCTCTTATG	CATGCTATACATGTTTGTGGCTATGAGGGCTATACACCCCGCTCC	1235	
Qy				
3141	CCACTCTCATAGCTG	CGCCGCGCTGGTTTGTGTCCAATGATGTACCATCATGCTG---	3197	
Db				
1236	TTATGCTATAGTGT	GATGATAGCTTATAGCCCTATAGGTGTGGGTATTATGACCATATTGAC	1295	
Qy				
3198	-----CTG	TGTCCTCCCTCTCGGGGCCCATGAGCGACTCCAAATTTCGAAAGGATCGAC	3249	
Db				
1296	CAC	TCCCCTATGTGTGAGATACATTTTCCATTATCTAATTCATAGATGGCTCTTTGGCCACA	1355	
Qy				
3250	ACCACGCTCACCTG	---CGAGACCCACCGTATCAACTGGACCGCCGATGACGACCTTGC	3306	
Db				
1356	ACTATCTATTGGCT	TATATGCCAATACTCTGTCCCTTCAGAGACTGACCGGACTCTGTA	1415	
Qy				
3307	GGCCTCAATGGAA	CGTTGTTCCTTCGACTGCTGCTCCGAGACCCGCCCCCA-----	3359	
Db				
1416	TTTTTTACAGAT	TGGGTCCATTTATTATTATACAAATTCACATATACAAACGCCGCTCC	1475	
Qy				
3360	-----AG	GCCCCGAGCACTCTGGATCAACTGCCCTTCGGCGGTTCGCGCTC	3409	
Db				
1476	CCGTGCCCGCAG	TTTTTATTAACAATAGCGTGGATCTCCGACATCTCGGGTACGTGTTT	1535	
Qy				
3410	AGCCCGGCGCG	TTTCACTTTCCCTTCAGAGCGGTCCCC-----CTTCAGCCCTACCA	3464	
Db				
1536	CGGACATGGCTCT	TTCGGTAGCGGGGAGCTTCACATCCGAGCCCTGGTCCCATCC	1595	
Qy				
3465	ATGCCAAT	TGCCCTCGGCTCTAGCGACGTTGCCCAATTATCGGGCAGCGCTTCTTCC	3524	
Db				
1596	GTCCAGCGGCT	CATGTCGGCTCGGAGCTCCTTGCTCCTAACTAGTGAGGCCAGACTTAG	1655	
Qy				
3525	CTGGAA	CAACTTAGTAAGGCATCTGTCTCTCAGAAAGTCTTATATTAATCAATGGC	3584	
Db				
1656	GCACAGCA	CAATGCCCAACCAACCAAGTGTGCCGACAAAGGCCGTGGCGGTATGT	1715	
Qy				
3585	CAATTTTCTCT	TACTCCCTCTTCGATACCTCTCTGTGGACCCCTCTCGGCTGCTCGT	3644	
Db				
1716	GTCTGAAA	TGAGTCTCGGATTTGGGCTCGCACTCGACGCGAGATGGAAGACTTAAGCA	1775	
Qy				
3645	CTTTGCC	CCCCAGACACCAAGGGAGCAATACGTTATCTCTCCACCCTTTTGACGTATGCC	3704	
Db				
1776	CGCGCAGA	AGAAGTGCAGGCGAGTGTGTTGTATTTCTGATAAGAGTCAGAGGTAAC	1835	
Qy				
3705	GGCTACTTG	TATTTCTA	CGCTAGCGAGCCCTCTCTCTTAATGTCCCATATGCCGCTT	3764
Db				
1836	CCCGTTCGG	TGTGTTTAACGGTGGAGGGCAGTGTAGTCTGACGAGTACTCGTGTCTGCC	1895	
Qy				
3765	TCCCGG	GACTCCAATGAACCCCTTTTCAGAAATTCGAGCTGCCCTTATCCAAAGCC	3824	
Db				
1896	CGCGCG	CCACAGACATAATAGCTGACAGACTA--ACAGACTGTGTTCTTCCATGGGTC	1953	
Qy				
3825	CGGCTGTCT	TGGTCTGTGCCCGCGATCGACCTATTCTCTAAACGGTCCCCCTTCCCATG	3884	
Db				
1954	TTTTCTG	CAGTCA	CGGTCTGCGAGTAATCAAGCAATCATGGATGCAATCAAGAGGGC	2013
Qy				
3885	CGACCG	GGTTACAGTATGGTCCATGCTCAGGCCCTTACAGCGCTTCTCTCATGACCCCTAC	3944	
Db				
2014	TCTGCTGT	GCTGCTGTGTGGAGCAGTCTTCGTTCGCCACGCGCTAGCGAAACCC	2073	
Qy				
3945	GCTAAC	CTGGTCCGAATAGTTGTCTAGCAGAAAAATAAGACTTGATCCCCCTTAAAT	4004	
Db				
2074	ACGTCA	CCGGGGAGGTGCCGCGC--CACATGTGTCTGGATTGTGTAGCTCTCTCGCACCA	2132	
Qy				
4005	ACAACT	GTCTAGAAAATGAATGGCTCTCCCGCTTTTTTTTGAGACCCAGCTTTCTTGACAA	4064	
Db				
2133	GGCGCC	AGCAGACGCTCCAGCTGATCAACACCAACGGCAGTGGCACCTCAATAGCACG	2192	
Qy				
4065	AGTGGT	GTATAACATCTGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGT	4124	
Db				
2193	GCCCTG	AACTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCAACCAC	2252	
Qy				

Db	4125	ATTCTTAACTATAGTTGGCTCTCTTTTACGCTATATGAGATACGCTGCTTTAATATGCTCTTTATGCTCTTTGTTAT	41184
Qy	2253	AAAGTTCAACTCTTTACAGGCTGTCTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGGATTTT	23112
Db	4185	CATGCTATTGCTTTCCCGTATGGCTTTCAITTTCTCTCCT-TGTATAAATCCTGGTTGCT	4243
Qy	2313	GACACGGCTTGGGGCCCTATCAGTTATGCGAACGGAAGCGGCCCGACACAGCGCCCTTAC	2372
Db	4244	GTCTCTTTATGAGGAGTTGTGGCCCTTGTCAGGCAACGTTGGCGTGGTGCTGTGCACTGTGTT	4303
Qy	2373	TGCTGSCATACCCCCCAAAACCTTGGGTATTTGTCCCGCGAAGAGTGTGTGTGTCG	2432
Db	4304	TGCTGACGCAACCCCC--ACTGGTTGGGCAITTTGCCACCACTGTGAGCTCTCTTTCGGG	4361
Qy	2433	GTATATTGCTTACTCCAGCCCCGTGGTGGGAACGACCGACAGGTGGGCGCGCCCC	2492
Db	4362	ACTTTGCTTTCCCTCTCCCTATTGCCAGCGGGAACCTATCGCGCCTGCCTTGGCCGC	4421
Qy	2493	ACCTACAGCTTGGGGTGAATAATGATACGGAAGTCTTCGTCTTTAAACAATACAGGCCACCG	2552
Db	4422	TGCTGGACAGGGG-----CTCGGCTTGGGCACTGACAATTCGTTGGTGTG	4469
Qy	2553	CTGGGCAATTGGTTTCGGTTGTACTTGGATGAACCTAACTGGATTACCAAAAGTGTGGGA	2612
Db	4470	TCGGGGAAATCATCGTCTCTTCTTGGCTGCTCGCCTGTGTGGCCACCTGGATTCTGCGC	4529
Qy	2613	GGCCCTCTTTGTCTCATCGAGGGGGGCAACACACACCTGCACGTGCCCACTGATTCG	2672
Db	4530	GGAGCTCTTCTGTCTACGTCCCTTCGGCCCTCAATCGAGCGGACCTTCTTCCCGGGC	4589
Qy	2673	TTCCGAAGCATCCGACGCCACATCTCTGGTGGGCTCGGTCCTCGATTCAGTCAACCC	2732
Db	4590	CTGCTCGGCTCTCGGCTCTTCCGCTCTTCGCTTCGCCCTCAGACGAGTCGGATC	4649
Qy	2733	AGTGCCCTGCTCGACTACCGGTATAGGCTTTGGCATATCTTTGTACCATCACTACACC	2792
Db	4650	TCCTTTGGGCCCTCTCCCGCTGTATCGATAAAATAAAAGATTTTATTAGTCTCCAGA	4709
Qy	2793	ATATTTAAAAATCAGGATGTACGTGGGAGGGTTCGAACACAGGCTGGAAGTGCCTGCAAC	2852
Db	4710	AAAAGGGGGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAAGTAACGCCAT	4769
Qy	2853	TGGACCGGGGCGAAGTTTCGATCTGGAAGATAGGACAGGTCGAGATCGATATGGAG	2912
Db	4770	TTTGCAAGCATCGAAAAATACATACTGAGAATACAGAAGTTTCAGATCAAGTCAAGAA	4829
Qy	2913	AACATACATCAGGATTCCTAGACCCCTGCTCGTGTTACAGCGGGGTTTTTCTTGTG	2972
Db	4830	CAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTCTCGTCCCGC	4889
Qy	2973	ACAAGAATCCTACAATACCGCAGAGTCTAGACTCGTGTGGTGGACTTCTCTCAATTTTCTA	3032
Db	4890	GCTACGGGCCAAGAACAGATAGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAA	4949
Qy	3033	GGGGGATCTCCCTGTGTCTTGGCCAAAAATTCGACGTCCCAACCTCCAATCACTCACCA	3092
Db	4950	GCAGTTCTGCCCGGCTCAGGGCCAAAGACAGATGGTCCCCAGATGC-----	4997
Qy	3093	ACCTCTGCTCTCCAAATTTGTCTCGTTATPCGCTGGATGTGTCTGGCGGTTTTATCAATA	3152
Db	4998	-GGTCCAGCCCTCAGCAGTTTCTTAGAGAACCATCAGATGTTTCCAGGGTG-----	5046
Qy	3153	TTCTCTTATCTGTCTGCTATGCCCTCATCTTCTTATTTGTTCTTCTGGATATCAAGT	3212
Db	5047	-----CCCCAAGACCTGAAATGACCCTGTGCTTTATTTGAACATAACCAATCAGTTTCGCT	5101
Qy	3213	ATGTTGCCCTTTGTCTCTTAATTCCAGGATCAACAAACACAGTACGGGACCATGCAAA	3272
Db	5102	TCTCGCTTCTGTTCCG-----GC	5119
Qy	3273	ACCTGCACGACTCTGCTCAAGGCAACTCTATGTTTTCTCTCATGTTGCTGTACAACCT	3332
Db	5120	GCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCTACTCGGGCGGCAGTCT	5179

Db	7284	GA	AACGTTCTT	CGGGCGGAAAACTCT	CAAGGATCTT	TACCGCTGTT	GAGATCCAGTTCGA	7343	
Qy	5550	TG	TAAACCCACT	CGTGCACCCCAACT	GTATCTTT	CAGCATCTTTT	TACTTTT	CACAGCGTTTCTG	5609
Db	7344	TG	TAAACCCACT	CGTGCACCCCAACT	GTATCTTT	CAGCATCTTTT	TACTTTT	CACAGCGTTTCTG	7403
Qy	5610	GG	TGAGCAAAAA	CAGGAAGGCAAAA	TGCCCAAAAAA	AGGGAATAA	AGGGCGCACACGGAAAT	5669	
Db	7404	GG	TGAGCAAAAA	CAGGAAGGCAAAA	TGCCCAAAAAA	AGGGAATAA	AGGGCGCACACGGAAAT	7463	
Qy	5670	GT	TGAATACTCA	TACTCTTCTCTT	TCTTCAATATAT	TATTTGAAGCATTTAT	CAGGGTTATGTCT	5729	
Db	7464	GT	TGAATACTCA	TACTCTTCTTCTT	TCTTCAATATAT	TATTTGAAGCATTTAT	CAGGGTTATGTCT	7523	
Qy	5730	TC	TATGAGCGGAT	CACATATTTT	GAATGTATTT	TAGAAAAAT	TAAACAATAGGGGTTCCCGCA	5789	
Db	7524	TC	TATGAGCGGAT	CACATATTTT	GAATGTATTT	TAGAAAAAT	TAAACAATAGGGGTTCCCGCA	7583	
Qy	5790	CA	TTTCCCCGAAA	AGTGCCACCTGAC	CTCTAAGAAAAC	CAATTTAT	TATCATGACATTAACCT	5849	
Db	7584	CA	TTTCCCCGAAA	AGTGCCACCTGAC	CTCTAAGAAAAC	CAATTTAT	TATCATGACATTAACCT	7643	
Qy	5850	AT	AAAAATAGGCGT	ATCACGAGGCCCTT	TCGTC	5882			
Db	7644	AT	AAAAATAGGCGT	ATCACGAGGCCCTT	TCGTC	7676			
RESULT 8									
US-10-401-000-2									
; Sequence 2, Application US/10401000									
; Publication No. US20040038304A1									
; GENERAL INFORMATION:									
; APPLICANT: Brenel, Robert D.									
; APPLICANT: Eakle, Kurt									
; APPLICANT: Imboden, Michael									
; TITLE OF INVENTION: Antibody Libraries									
; FILE REFERENCE: GALA-07886									
; CURRENT APPLICATION NUMBER: US/10/401,000									
; CURRENT FILING DATE: 2003-03-27									
; PRIOR APPLICATION NUMBER: 60/368,808									
; PRIOR FILING DATE: 2002-03-28									
; PRIOR APPLICATION NUMBER: 60/371,299									
; PRIOR FILING DATE: 2002-04-10									
; NUMBER OF SEQ ID NOS: 4									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2									
; LENGTH: 7626									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Synthetic									
US-10-401-000-2									
Query Match 39.5%; Score 2321.4; DB 17; Length 7626;									
Best Local Similarity 67.4%; Pred. No. 0;									
Matches 3744; Conservative 0; Mismatches 1566; Indels 243; Gaps 24;									
Qy	337	ATT	TATGGCTATTGGCC	CAATGCGTACGTTGTATCTATAT	CATATATATGTCATATTTATTT	396			
Db	2301	AAT	TATGGCTATTGGCC	CAATGCGTACGTTGTATCTATAT	CATATATATGTCATATTTATTT	2360			
Qy	397	GG	CTCATGTC	CCATATGACGCCCATGTTG	CATTTGATTTTGTACTAGTTATTAATAGTAA	456			
Db	2361	GG	CTCATGTC						

Qy	1567	C A G A C A C A A T T G C C C C A C C A C C A C A G A C T G T G C C G C A C A A G C C G T G G C G G T A G G T A T G T G	1711
Db	3606	A G C A C C T C T G G G G C C A C A G C G G C C T T G G G C T G C C -----	3639
Qy	1717	T C T G A A A T A G A G C T C G A G A T T G G G C T C G C A C C T T G G A C G C A G A T G G A A G A C T T A A G C G A G	1776
Db	3640	-----T G G T C A A G A C T A C T T C C C C G A A C C G G T G A C G G T G C T G T G A A C T C A G ---G	3688
Qy	1777	C G G C A G A A G A G T G C A G G C A G C T G A G T T G T T G T A T T C T G A T A A G A G T C A G A G G T A A C T C	1836
Db	3689	G C C C T A C C A G C G G C G T C A C A C C T T C C C G G C T G C C T A C A G T C C T C A G S A C T C T A C T C	3748
Qy	1837	C G T T T C G G T G C T G T T A A C G G T G G A G G C A G T G A G T C T G A C A G A G T A C T C G T T G C T G C C G	1896
Db	3749	C C T C A G C A G C G T G T G T A C C G T G C C T C C A G C A G C T T G G G C A C C C A G A C C T A C A T C T C G C A A	3808
Qy	1897	C G G C G C C A C C A G A C A T A A T A G C T G A C A G A C T A A C A G A C T G T T C C T T T C C A T G G G T C T T T	1956
Db	3809	C G T G A T C A A A G C C C A G A A C A C C A A G G T G G A C A A G A G A T T - G A G C C C A A A T C T T T G T G	3867
Qy	1957	T C T G C A G T C A C C G T C G T C G A C G A A T T C A A G C A A T C A T G S A T G C A A T G A A G A G A G G G C T C T	2016
Db	3868	A C A A A C T C A C A C A T G C C C A C C G T G C C C A G C A -----	3899
Qy	2017	G C T G T G T G C T G C T G T G T G A G A C A G T C T T C G T T T C G C C C A G C G G T A G C G A A A C C C A C G	2076
Db	3900	--C C T G A A C T C T G G G G G A C C G G T C A G T C T T C C T C T T C C C C -----C C A A A A C C C A A G G	3951
Qy	2077	T C A C G G G G N A G T G C C G G C C A C A C T G T G T C T G G A T T T G T A G C C T C C T C G C A C C A G G G	2136
Db	3952	A C A C C T C A T G A T C T C C C G A C C C C T A G A G T C A C A T G C G T G G T G G A C G T A G C C A C G	4011
Qy	2137	C C A A G C A G A A C G T C C A C G C T G A T C A A C A C A A C G C A G T T G C G C A C C T C A A T A G C A C C G C C C	2196
Db	4012	A A G A C C T G A G T C A A G T T C A A C T G G T A C G T G A C G C G T G A G G T G C A T A T G C C A A G A	4071
Qy	2197	T G A A C T C A A T A G A P A G C C T C A A C A C C G G T G G T T G G C A G G G C T T T C T A T C A C A A A G T	2256
Db	4072	C A A A G C C G G G A G A G C A G T A C A A C A G C A C G T A C C G T G T G G T C A G C G T C C T C A C C G T C C	4131
Qy	2257	T - C A A C T T C A G G C T G T C T G A G A G C T A G C C A G C T G C C G A C C C T T A C C G A T T T T G A C	2315
Db	4132	T G C A C C A G A C T G G C T G A A T G G C A A G A G A T A C A A G T G C A A G G T C T C C A A C A A A G C C C T C C	4191
Qy	2316	C A G G C T G G G G C C C T A C A G T T A T G C C A A C G A A G C G C C C C A C A G C C C C C T A C T C G C	2375
Db	4192	C A G C C C C A T C G A A A A C C A T C T C A A A G C C A A A G G C A G C C C C G A A C C A C A G T G T	4251
Qy	2376	T G G C A C T A C C C C C A A A C C T T G C G G --T A T T G T G C C C G A A G A G T G T G T G G T C C G G	2433
Db	4252	A C A C C T G C C C C A T C C C G G A G A G A T A C C A A G A A C A G G C A G C C C G A A C C A C A G T G T	4311
Qy	2434	T A T A T T G C T T C A C T C C C A G C C C G T G T G T G G G A C C G A C A G T C G G G C G C G C C C A	2493
Db	4312	T C A A A G G C T T C T A T C C C A G C G A C A T C C C G T -----G G A G T G G G A G A G C A A	4357
Qy	2494	C C T A C A G C T G G G T G A A A T G A T A C G A C G T C T T T G C T C T T A C A A T A C C A G C C C A C C G C	2553
Db	4358	T G G C A C C C G G A A C A A C T A C A A G A C C A C G C C T C C G T G C T G A C T C C A C G G C T C T T	4417
Qy	2554	T G G G C A A T T G G T T T G T T A C T C T G A T A A C T C A A C T G G A T T C A C C A A A G T G T G G G A G	2613
Db	4418	C T T C C T A T A G A A G C T C A C C G T G G A C A A G A G C A G G T G G C A G A -----G G G A A	4468
Qy	2614	G C G C T C C T T G T G T C A T C G G A G G C G G G C A A C A A C C C T G C A C T G C C C A C T G A T T G C T	2673
Db	4469	C G T C T T C A T G C T C G T G A T G A C A G A G G C T C T G C A A C C A C C A C T A C A C G C A A A G A G C C T	4528
Qy	2674	T C G C A A G C A T C C G A C C C A C A T A C T C T C G G T G C G C T C C G G T C C C T G A T C - A C A C C C	2732
Db	4529	C T C C T G T C C C G G G A A A T A A A A G C C G A A T T C C G G C C C A C T G A G A T A T C T A G A C C C	4588
Qy	2733	A G G T G C C T G C A C T A C C C G T A T A G C C T T T G G C A T T A T C C T T G T A C A A T C A A C T A C A C C	2792

Db	4589	AGCTTTCTTGTCACAAAGTGGTGATA-----ACATCGATAAAATAAAGATT	4634
Qy	2793	ATATTTAAAAATCAGGATGTACGTGGGAGGGTCCGAACACAGAGCTGGAGCTGCCTGCAAC	2852
Db	4635	TTATTTAGTCTCCAGAAAAAGGGGGGATCAAGACCCCA-CTGTAGGTTTGGCAAGCT	4693
Qy	2853	TGGACGGGGCGAAACGTTCCGATCTTGGAAAGATAGGGACAGGTCGGAGATCGATATCGAG	2912
Db	4694	AGCTTAAGTAACGGCCATTTTTCGAAGGCATGCAAAAAATACATAACTGAGA-----ATAGAG	4748
Qy	2913	AACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTATACAGGGGGGTTTTTCTGTGTTG	2972
Db	4749	AAGTTCAAGTCAAGGT-----CAGGAACAGATGGGAACAGCTG	4785
Qy	2973	ACAAGAATCTCTACAATACCGCAGAGTCTAGACTCGTGGTGGATCTCTCTCAATTTTCTTA	3032
Db	4786	AATATGGGCCAAACAGGATATCTGTGTAGACAGTTCTCCTCCCGGCTCAGGGCCAGAA	4845
Qy	3033	GGGGGATCTCCCGTGTGTTCTTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCAACA	3092
Db	4846	CAGATGAAACAGCTGAAATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCCG	4905
Qy	3093	ACCTCTGTCTTCCAATTTGTCTGTGTTATCGCTGGATGTGCTGGCGCGTTTATATCATTA	3152
Db	4906	GCTCAGGGCCAAAGAACAGATGGTCCCCAGATGGGTCACGCCCTCAGCAGTTTCTTAGAGA	4965
Qy	3153	TTCTCTCTTCATCTGCTGCTATGCCTCATCTTCTTATTGGTCTCTCTCGAATTTATCAAGGT	3212
Db	4966	ACATCAGATGTTTCCAGGGTGCCCCAGACCTGAATGACCTGTGCTTATTTTGAAC	5025
Qy	3213	ATGTGGCCCGTTGTCTCTTAATTCAGGATCAACAACAACAGTAGTAGGGACCATGCAAA	3272
Db	5026	TAACCAATCAGTTTCGCTTCTCGTTCGTTCGGCGGCTTCTGTCTCCCGAGCTCAATAAA	5085
Qy	3273	ACCTGCAGCATCTCTGCTCAAGGCAACTATGTTTCCCTCATGTGTTCTGTATCAAAACCT	3332
Db	5086	AGAGCCCAACAACCCCTCACTCGGGGGGCCAGTCTCTCGATTTGATCGAGTCGCCGGGTAC	5145
Qy	3333	ACGGATGGAATATGCACCTGATTTCCCATCCCATCGTCGGGCTTTCGCAAAATACCTA	3392
Db	5146	CCGTGATCAATAAACCCCTCTTGAGTTGCATCCGACTTGTGGTCTCGC-----5195	
Qy	3393	TGGGAGTGGGCCTCAGTCCGTTTTCTCTTGGCTCAGTTTACTAGTGCATTTGTTCAGTGG	3452
Db	5196	-----TGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCG	5235
Qy	3453	TTCGTAGGGCTTTCGCCCACTGTTTGGCTTTACGCTATATATGATATGTGGTATTTGGGGG	3512
Db	5236	TCAGCGGG-----GTCTTTTCAATTTTTCATTTGGGGGCTCGTCCGGGATCGGG	5283
Qy	3513	CCAAGTCTGTACAGCATCGTGGTCCCTTTATACCGCTGTATACCAATTTTCTTTT---GT	3569
Db	5284	AGACCCCTGCCAGGGACCAACGACCCCAACCGGGAGGTAAAGCTGGCTGCCCTCGCGCGT	5343
Qy	3570	CTCTGGGTATACATTTAAGAAATTCAGACTCGACAGTCTAGAAAAGCGCGCCAAAGATAT	3629
Db	5344	TTCCGTGATCGCGGTGAAACCTCTGNACATATGAGCTCCCGGAGCGGTCAAGTTGT	5403
Qy	3630	CAAGGATCCACTACGGGTTAGAGCTCGCTGATCAGCCTCGACTGTGTGCTTCTATGTGCCA	3689
Db	5404	CTGTAAAGCGGATGCCGGGAGCAG-----ACAAGCCCGTCAGGGCGC	5444
Qy	3690	GCCATCTGTTGTTTGCCCTTCCCGCTCCCTTTCGACCTGGAGGTGCCACTCCAC	3749
Db	5445	GTACAGCGGTTGTTGGCGGGTGTCCGGGGCGAGCCATGACCCAGTCAAGTATAGCATAGCGG	5504
Qy	3750	TGTCCTTTCTTAATAAATGAGGAAATTCATCGCATTTGTCTGAGTAGGTGTGATTCTAT	3809
Db	5505	AGTGAT-----ACTGGCTTAACTATGGGCATCAGACAGATTGTTAC	5547
Qy	3810	TTCTGGGGGTGGGTGGGGCAGGACAAAGGGGGAGGATTTGGGAAGACAATAGACGCCA	3869

Db 5548 TGAGAGTGCACCATATATCGCGTGTGAATACCGCACAGATCGCTAAGGAGAAATACCGCA 5607
QY 3870 TCGTGGGAGCTCTTCGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGC 3929
Db 5608 ---TCAGGGCGCTCTTCGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTTCGGCTGC 5664
QY 3930 GCGAGCGGTATACAGTCACTCAAGCGCGTAATACGGTTATCCACAGAAATCAGGGGATA 3989
Db 5665 GCGAGCGGTATACAGTCACTCAAGCGCGTAATACGGTTATCCACAGAAATCAGGGGATA 5724
QY 3990 ACCAGAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCCGTAAGGAGCGG 4049
Db 5725 ACCAGAGAAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCCGTAAGGAGCGG 5784
QY 4050 CGTTGCTGGCGTTTTCATPAGGCTCCGCCCTCGACGAGCATCAAAAAATFCGACGCT 4109
Db 5785 CGTTGCTGGCGTTTTCATPAGGCTCCGCCCTCGACGAGCATCAAAAAATFCGACGCT 5844
QY 4110 CAAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAAGGCGTTTCCCTCGAA 4169
Db 5845 CAAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACCAAGGCGTTTCCCTCGAA 5904
QY 4170 GCTCCCTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTC 4229
Db 5905 GCTCCCTCGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTC 5964
QY 4230 TCCCTTGGGAGCGTGGCGCTTTCATATGCTCACGCTGTAGGTATCTCAGTTCGGTGT 4289
Db 5965 TCCCTTGGGAGCGTGGCGCTTTCATATGCTCACGCTGTAGGTATCTCAGTTCGGTGT 6024
QY 4290 AGGTGCTTCGCTCAAGCTGGGCTGTGCGAAGACCCCGCTTACGCGCGACCGCTCG 4349
Db 6025 AGGTGCTTCGCTCAAGCTGGGCTGTGCGAAGACCCCGCTTACGCGCGACCGCTCG 6084
QY 4350 CTTTATCCGCTAACTATCGTCTTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTG 4409
Db 6085 CTTTATCCGCTAACTATCGTCTTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTG 6144
QY 4410 CAGACGCCACTGGTAAACAGATTAGCAGAGCGAGGTATGAGCGGTGTACAGAGTTCT 4469
Db 6145 CAGACGCCACTGGTAAACAGATTAGCAGAGCGAGGTATGAGCGGTGTACAGAGTTCT 6204
QY 4470 TGAAGTGGTGGCTTAAGTCACTAGAGGACAGTATTTGGTATCTCGCTCGC 4529
Db 6205 TGAAGTGGTGGCTTAAGTCACTAGAGGACAGTATTTGGTATCTCGCTCGC 6264
QY 4530 TGAAGCCAGTTACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCCG 4589
Db 6265 TGAAGCCAGTTACCTTTCGGAAGAGTTGGTAGCTCTTGATCCGCGCAACCAACCCG 6324
QY 4590 CTGGTACGCGTGGTTTTTTTGGTTGCAAGCAGAGATTACGCGCAAGAAAGGATCTC 4649
Db 6325 CTGGTACGCGTGGTTTTTTTGGTTGCAAGCAGAGATTACGCGCAAGAAAGGATCTC 6384
QY 4650 AAGAAGTCTTTGATCTTTCTACCGGCTCGACCTCAGTGGACGAAACTCAGCTT 4709
Db 6385 AAGAAGTCTTTGATCTTTCTACCGGCTCGACCTCAGTGGACGAAACTCAGCTT 6444
QY 4710 AAGGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAA 4769
Db 6445 AAGGATTTTGGTCATGAGATTATCAAAAGGATCTTCACTAGATCTTTTAAATTTAA 6504
QY 4770 AATGAAGTTTAAATCAATCAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 4829
Db 6505 AATGAAGTTTAAATCAATCAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT 6564
QY 4830 GCTTATCAGTGGGACCACTATCTCAGCATCTGCTATTTCTTCTCATCATAGTTGCT 4889
Db 6565 GCTTATCAGTGGGACCACTATCTCAGCATCTGCTATTTCTTCTCATCATAGTTGCT 6624
QY 4890 GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTG 4949
Db 6625 GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGCTTACCATCTGGCCCCAGTGTG 6684

QY 4950 CAAATGATACCGCAGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCAG 5009
Db 6685 CAAATGATACCGCAGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAAACCCAGCAG 6744
QY 5010 CCGGAAGGCGCCAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 5069
Db 6745 CCGGAAGGCGCCAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 6804
QY 5070 ATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCCAAGTTAATAGTTTGCACACGTTGTTG 5129
Db 6805 ATTGTTGCGGGAAGCTAGAGTAAGTAGTTCCCAAGTTAATAGTTTGCACACGTTGTTG 6864
QY 5130 CCAATTGCTACAGGCATCGTGTGTCAGCTCTGTTGGTATGCTTCATTTCAGTCTCG 5189
Db 6865 CCAATTGCTACAGGCATCGTGTGTCAGCTCTGTTGGTATGCTTCATTTCAGTCTCG 6924
QY 5190 GTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGCT 5249
Db 6925 GTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGTTAGCT 6984
QY 5250 CCTTCGGTCTCCGATCGTTGTGAGAAAGTTTGGCCGAGTGTATCACTCATGGTTA 5309
Db 6985 CCTTCGGTCTCCGATCGTTGTGAGAAAGTTTGGCCGAGTGTATCACTCATGGTTA 7044
QY 5310 TGGCAGCATGCTAATTTCTTACTGTATGCCCATCGTAAGATGCTTTCTGTGACTG 5369
Db 7045 TGGCAGCATGCTAATTTCTTACTGTATGCCCATCGTAAGATGCTTTCTGTGACTG 7104
QY 5370 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGCGACCGAGTGTCTTGGC 5429
Db 7105 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGCGACCGAGTGTCTTGGC 7164
QY 5430 CGGCGTCAATACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCTG 5489
Db 7165 CGGCGTCAATACGGGATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCTG 7224
QY 5490 GAAAAGTCTTTCGGGGGCAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTCCA 5549
Db 7225 GAAAAGTCTTTCGGGGGCAAACTCTCAGGATCTTACCGCTGTGAGATCCAGTTCCA 7284
QY 5550 TGTAAACCCACTCGTGCACCCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTCTG 5609
Db 7285 TGTAAACCCACTCGTGCACCCAACTGATCTTACGATCTTTTACTTTTACCAGCGTTTCTG 7344
QY 5610 GGTGAGCAAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGAAT 5669
Db 7345 GGTGAGCAAAAAACAGGAAGGCAAAATGCGCAAAAAAGGAATAAGGGCGACACGGAAT 7404
QY 5670 GTTGAATACTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5729
Db 7405 GTTGAATACTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7464
QY 5730 TCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAAATAGGGGTTCCCGCA 5789
Db 7465 TCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAAATAGGGGTTCCCGCA 7524
QY 5790 CATTTTCCCGAAAAAGTGCCACTGACGTCTAAGAAACCAATTTATTCATGACATTAACCT 5849
Db 7525 CATTTTCCCGAAAAAGTGCCACTGACGTCTAAGAAACCAATTTATTCATGACATTAACCT 7584
QY 5850 ATAAAAATAGGCGTATCAGAGCGCTTTCGTC 5882
Db 7585 ATAAAAATAGGCGTATCAGAGCGCTTTCGTC 7617

RESULT 9

US-10-401-000-1
; Sequence 1, Application US/10401000
; Publication NO. US2004003804A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert D.
; APPLICANT: Bakle, Kurt

APPLICANT: Imboden, Michael
; TITLE OF INVENTION: Antibody Libraries
; FILE REFERENCE: GALA-07886
; CURRENT APPLICATION NUMBER: US/10/401,000
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: 60/368,808
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/371,299
; PRIOR FILING DATE: 2002-04-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 7617
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-401-000-1

Query Match 39.3%; Score 2311.8; DB 17; Length 7617;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 3734; Conservative 0; Mismatches 1567; Indels 252; Gaps 23;

QY	337	ATTATTGGCTATTGGCCATTGGCATAGCTTGTATCTATATCATTAATATGATACATTATATT	396
DB	2301	AATATTGGCTATTGGCCATTGGCATAGCTTGTATCTATATCATTAATATGATACATTATATT	2360
QY	397	GGCTCATGTCCATATAGCCGCATGTGACATTGATTATGACTAGTTATTAATAGTAA	456
DB	2361	GGCTCATGTCCATATAGCCGCATGTGACATTGATTATGACTAGTTATTAATAGTAA	2420
QY	457	TCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCGCGGTACATACTTACG	516
DB	2421	TCAATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCGCGGTACATACTTACG	2480
QY	517	GTAATAGGCCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCATTAATGACG	576
DB	2481	GTAATAGGCCGCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCATTAATGACG	2540
QY	577	TATGTTCCCATAGTAAGCCCAATAGGACTTTCATAGCTCAATGGGTGGAGTATTTA	636
DB	2541	TATGTTCCCATAGTAAGCCCAATAGGACTTTCATAGCTCAATGGGTGGAGTATTTA	2600
QY	637	CGGTAACATGCCCACTTGGCAGTACATCAAGTGTATCATATGATGCAAGTCCGCCCTATT	696
DB	2601	CGGTAACATGCCCACTTGGCAGTACATCAAGTGTATCATATGATGCAAGTCCGCCCTATT	2660
QY	697	GACGTCAATGACGGTAAATGGCCGCTGGCATTATGCCAGTACATGACCTTACGGGAC	756
DB	2661	GACGTCAATGACGGTAAATGGCCGCTGGCATTATGCCAGTACATGACCTTATGGGAC	2720
QY	757	TTTCTTACTTGGCAGTACATCTAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT	816
DB	2721	TTTCTTACTTGGCAGTACATCTAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT	2780
QY	817	TGGCAGTACACCAATGGCGTGATAGCGGTTTGACTCAGCGGGATTTTCCAACTCTCCAC	876
DB	2781	TGGCAGTACACCAATGGCGTGATAGCGGTTTGACTCAGCGGGATTTTCCAACTCTCCAC	2840
QY	877	CCATTGACGTCAATGGGAGTTGTTTGTGGCACCAAAATCAACGGGACTTTTCCAAATGT	936
DB	2841	CCATTGACGTCAATGGGAGTTGTTTGTGGCACCAAAATCAACGGGACTTTTCCAAATGT	2900
QY	937	CGTAATAACCCCGCTTGGACCAATGGGCGGTAGGCGGTATAGTGGGAGGTCTAT	996
DB	2901	CGTAATAACCCCGCTTGGACCAATGGGCGGTAGGCGGTATAGTGGGAGGTCTAT	2960
QY	997	ATAAGCAGAGCTCGTTTATGTAACCGTCAGATCGCTGGAGACGCCATCCAGCGCTGTTT	1056
DB	2961	ATAAGCAGAGCTCGTTTATGTAACCGTCAGATCGCTGGAGACGCCATCCAGCGCTGTTT	3020
QY	1057	GACCTCCATAGAGACACCGGGACCGATCCAGCTCCCGCGCGGGAAACGGTGCAATTGGA	1116

DB	3021	GACCTCCATAGAGACACACGGGACCGATCCAGCTCCCGGGCCCCAAGCTTGTATTATCACA	3080
QY	1117	ACGGGATTTCCCGTCCCAAGAGTAGTACGTAAGTACCGCTATATAGACTCTATAGGCACACC	1176
DB	3081	AGTTTGATACAAAAGCAGGCTTCGAGGAGATAGAACCAATTTCTTAAGGAATACTTA	3140
QY	1177	CTTTTGGCTCTTATGCAATGCTATATCTGTTTGGTGTGGGGCTATACACCCCGCTCT	1236
DB	3141	ACCATGGGATGGAGCTGTATCATCTCTTCTTGG-TAGCAACACAGCTACAGGTCTCCACTC	3199
QY	1237	TATGCTATAGGTGATGATAGCTTAGCTTATAGGTGTGGTATTGACCAATTATTGACC	1296
DB	3200	CCAGGTCCAGCTGGTCCAAAT--CAGGGGCTGAAGTCAAGAAACCTGGGTATCAGTGAAG	3257
QY	1297	ACTCCCTTATTTGGTACAGTACTTTTCCATTACTTAATCAATCAATGGCTCTTTGCCACAA	1356
DB	3258	GTCTCTCTCAAGGCTTCGGCTACACCTTTACTAGCTACTGGCTGACCTGGGTACGGCAG	3317
QY	1357	CTATCTTATTTGGCTATATGCCAATACTCTGCTCTTTCAGAGACTGACCGGACTCTGTAT	1416
DB	3318	GCACCTGGACAGGGTCT-----GGAATGGATTGGATACATTAATCTTAGGAATGA	3367
QY	1417	TTTTACAGATGGGGTCCATTTTATTATTACAAATTCATATATACAAACGCGCTCCCC	1476
DB	3368	TTATATCTGATACAATCAGAACTTCAAGGACCAAGGCCCAATAAATCTGCAGCAATCCAC	3427
QY	1477	CGTGGCCGAGTTTTTTTAAACATAGCGTGGGATCTCCGACATCTCGGTACGTGTTCC	1536
DB	3428	CAATACAGCTTACATGGAGCTGACGAGCTTCCACATCCGAGGCCCTGGTCCCATCCG	1596
QY	1537	GGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTCCACATCCGAGGCCCTGGTCCCATCCG	1596
DB	3488	TGCAA-----GAAGGATATTTACTAGCTTCTACTGGGGCCAAAGGC	3527
QY	1597	TCAGCGGCTCATGTTGGTTCGGCTCGGAGCTCTTGTCTCTTAACAGTGAGGCGCAGACTTAG	1656
DB	3528	ACCAGCTCACCGTCTCTCTCAGCTCCACAAAGGGCCATCGGTCTTCCCTCGGCACCC	3587
QY	1657	CACAGCACAAATGCCACACCAACCCAGTGTCCGACACAGGCCGCGGTGAGGTATGTG	1716
DB	3588	TCCTCAAGAGCACCTCTGGGGGACAGCGGCCCTGGGCTGCTGTCTAAGACTACTTTC	3647
QY	1717	TCGAAAATGAGCTCGAGATTGGGCTCGCACCTCGACAGCATGGACGAGTGAAGACTTTAAGGCAG	1776
DB	3648	CCGAAACCGTGACGGTGTCTGG-----AACTCAGG	3679
QY	1777	CGGAGAGAAAGATGACGAGCTGAGTTGTTGTTATTTCTGATTAAGAGTCAGAGTAACTC	1836
DB	3680	CGCCCTGACAGCGGGTGACACACCTTCCCGGCTGCTTACAGTCTCTCAGGACTCTACTC	3739
QY	1837	CGTTTGGGTGCTGTTTAAAGGTGGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTGCCG	1896
DB	3740	CCTCAGACGGTGTGACCGCTCCAGCAGCTTGGGGACCCAGACCTACATCTGCAA	3799
QY	1897	CGCGGCCACCAAGACATAATAGCTGACAGACTAAACAGACTGTTTCTTCCATGGGCTTTT	1956
DB	3800	CGTGAATCACAAGCCAGCAACACCAAGGTGGCAAGAGAGTTT-GAGCCCAATCTTGTG	3858
QY	1957	TCGCAAGTCAACGCTGCTGACAGAAATTAAGCAATCATGGATGCAATGAAGAGGGCTCT	2016
DB	3859	ACAAAACCTCACATGCCCCACCGTGCCACGA-----3890	
QY	2017	GCTGTGCTGCTGCTGTGTGGAGCAGTCTCTGTTTGGCCCGCTAGCGAAACCCACG	2076
DB	3891	--CCTGAACTCTCTGGGGGACCGCTCAGTCTTCTCTTCTTCCCT-----CCAAAACCCAAAGG	3942
QY	2077	TCACCGGGGAAAGTGGCGGCCACACTGTCTGTGATTGTTAGTCTCTCTCGCACCCAGGCG	2136
DB	3943	ACACCTCATGATCTCCCGGACCCCTGAGGTACATGCGTGGTGGTGGAGTGGAGCCACG	4002
QY	2137	CCAAGCAGAACTGCTGAGCTGATCAACCAACGCGCAGTTGGCACTCAATAGACCGGCC	2196
DB	4003	AAGACCTGAGGTCAAGTTCAACTGGTACGTGGACCGCGTGGAGTGCATTAATGCCAAGA	4062

Db 6076 CTTTATCCGGTAACATATCGTCTTGGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGG 6135
QY 4410 CAGCAGCAGTCTGGTAACAGGATAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT 4469
Db 6136 CAGCAGCAGTCTGGTAACAGGATAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT 6195
QY 4470 TGAAGTGGTGGCTTAACACTACGGCTACACTAGAAGGACAGATTTGGTATCTCGGCTCTGC 4529
Db 6196 TGAAGTGGTGGCTTAACACTACGGCTACACTAGAAGGACAGATTTGGTATCTCGGCTCTGC 6255
QY 4530 TGAAGCAGTCTCTCGGAAAGAGTTGGTAGTCTTGTATCGGCAAAACAAACACCG 4589
Db 6256 TGAAGCAGTCTCTCGGAAAGAGTTGGTAGTCTTGTATCGGCAAAACAAACACCG 6315
QY 4590 CTGCTAGCGGTGGTTTTTGTGTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCTC 4649
Db 6316 CTGCTAGCGGTGGTTTTTGTGTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCTC 6375
QY 4650 AAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTT 4709
Db 6376 AAGAAGATCCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCAGTT 6435
QY 4710 AAGGATTTTGGTCATGAGATTTACAAAAAGATCTTCACTAGATCTCTTTTAAATTTAA 4769
Db 6436 AAGGATTTTGGTCATGAGATTTACAAAAAGATCTTCACTAGATCTCTTTTAAATTTAA 6495
QY 4770 AATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACCTTGGTCTGACAGTTACCAT 4829
Db 6496 AATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACCTTGGTCTGACAGTTACCAT 6555
QY 4830 GCTTAATCAGTGGACCACTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCT 4889
Db 6556 GCTTAATCAGTGGACCACTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCT 6615
QY 4890 GACTCCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGGTG 4949
Db 6616 GACTCCCCGTGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGGTG 6675
QY 4950 CAATGATACCGCAGACCCAGCTCACCGGCTCCAGATTTATCAGCAATAAACAACCCAG 5009
Db 6676 CAATGATACCGCAGACCCAGCTCACCGGCTCCAGATTTATCAGCAATAAACAACCCAG 6735
QY 5010 CCGGAAGGGCGAGCGCAGAGTGGTCTCCTCAACTTTATCCGCTCCATCCAGTCTATTA 5069
Db 6736 CCGGAAGGGCGAGCGCAGAGTGGTCTCCTCAACTTTATCCGCTCCATCCAGTCTATTA 6795
QY 5070 ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAACGTTGTTG 5129
Db 6796 ATTGTTGCCGGGAAGCTAGAGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAACGTTGTTG 6855
QY 5130 CCATTGCTACAGGCATCGTGGTCTCAGCTCGTGGTTGGTATGGCTTCAATCAGCTCCG 5189
Db 6856 CCATTGCTACAGGCATCGTGGTCTCAGCTCGTGGTTGGTATGGCTTCAATCAGCTCCG 6915
QY 5190 GTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCT 5249
Db 6916 GTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCT 6975
QY 5250 CTTCCGGTCTCCGATCGTTGTGAGTAAGTTGGCCGAGTGTATCACTCATGTTA 5309
Db 6976 CTTCCGGTCTCCGATCGTTGTGAGTAAGTTGGCCGAGTGTATCACTCATGTTA 7035
QY 5310 TGGCAGCATGCAATAATCTTCTACTGTCATGTCATCCGTAAGATGCTTTCTGCACTG 5369
Db 7036 TGGCAGCATGCAATAATCTTCTACTGTCATGTCATCCGTAAGATGCTTTCTGCACTG 7095
QY 5370 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCACCGAGTTGCTTTGCC 5429
Db 7096 GTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCACCGAGTTGCTTTGCC 7155
QY 5430 CGCGGTCAATACCGGGAATAATACCGGCCCAATAGCAGAACTTTTAAAGTGTCTCATCTG 5489

Db 7156 CGCGGTCAACACACGAGTAATAATACCGGCCCATACAGCAAACTTTTAAAGTGTCTCATCTG 7215
QY 5490 GAAAAAGTTCTTCCGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCA 5549
Db 7216 GAAAAAGTTCTTCCGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCA 7275
QY 5550 TGTAAACCCACTCGTGACCCCAACTGATCTTTCAGCATCTTTTACTTTCACGAGGTTTCTG 5609
Db 7276 TGTAAACCCACTCGTGACCCCAACTGATCTTTCAGCATCTTTTACTTTCACGAGGTTTCTG 7335
QY 5610 GGTGAGCAAAAAACAGAGCAAAATGCCGCAAAAAAGGAATAAGGCGACACGGAAT 5669
Db 7336 GGTGAGCAAAAAACAGAGCAAAATGCCGCAAAAAAGGAATAAGGCGACACGGAAT 7395
QY 5670 GTTGAATACTCATACTCTTCTTCTTCAATATTTTCAATATTTTGAAGCATTTTATCAGGTTTATGTC 5729
Db 7396 GTTGAATACTCATACTCTTCTTCTTCAATATTTTGAAGCATTTTATCAGGTTTATGTC 7455
QY 5730 TCATGAGCGGATACATATTTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCA 5789
Db 7456 TCATGAGCGGATACATATTTTGAATGTTATTTAGAAAAATAAACAATAGGGGTTCCGCGCA 7515
QY 5790 CATTTCCCGAAAAAGTGCCACCTGACGCTTAAGAAACCATTTATATCATGACATTAACCT 5849
Db 7516 CATTTCCCGAAAAAGTGCCACCTGACGCTTAAGAAACCATTTATATCATGACATTAACCT 7575
QY 5850 ATAAATAAGGCGTATCAGGAGCCCTTTTCGTC 5882
Db 7576 ATAAATAAGGCGTATCAGGAGCCCTTTTCGTC 7608

RESULT 10

US-10-359-120-50
; Sequence 50, Application US/10359120
; Publication No. US20040033487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKRABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; TITLE OF INVENTION: ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 6577
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
US-10-359-120-50

Query Match 38.9%; Score 2290.6; DB 17; Length 6577;
Best Local Similarity 68.9%; Pred. No. 0;
Matches 3636; Conservative 0; Mismatches 1139; Indels 502; Gaps 16;
QY 1 TCGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
Db 1 TCGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAGCGGATGCCGAGCAGACAAGCCCGTCAGGCGGTG 120
Db 61 CAGCTTGTCTGTAAGCGGATGCCGAGCAGACAAGCCCGTCAGGCGGTG 120

QY	121	TTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATGTGTAAGAGTGC	180
DB	121		
QY	181	TTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATGTGTAAGAGTGC	180
DB	181		
QY	181	ACCATATGAAGCTTTTGTGMAAGCCTAGGCCTCCAAAAAGCCTCCTCACTACTTCTGG	240
DB	181		
QY	241	ACCATATGGG-----TGTGAAATACCGCAC	207
DB	241		
QY	208	AATAGCTCAGAGCGCAGGGCGGCTCGGCCTCTGCATAATAAAAAAATTTAGTCAGCCA	300
DB	208		
QY	301	GATGCGTAGGAGAAATACCGCATCAG-----	235
DB	301		
QY	236	TGGGCGGAGAAATGGGCGGAATGGGCGGGAGGGAATTTAGGCTATTGGCCATTGCAAT	360
DB	236		
QY	361	ATTTGGCTATTGGCCATTGCAAT-----ATTGGCTATTGGCCATTGCAAT	256
DB	361		
QY	361	ACGTTGATCTATCATATAATGTACATTTATATTGGCTCATGTCCCAATATGACCGCCA	420
DB	361		
QY	257	ACGTTGATCTATCATATAATGTACATTTATATTGGCTCATGTCCCAATATGACCGCCA	316
DB	257		
QY	421	TGTTGACATTTGATTTGACTAGTTTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT	480
DB	421		
QY	317	TGTTGACATTTGATTTGACTAGTTTAAATAGTAATCAATTACGGGGTCATTAGTTTCAT	376
DB	317		
QY	481	AGCCCATATATGAGTTTCGCGTTACATAACTTACGGTAAATGGCCGCGCTGACTGACCG	540
DB	481		
QY	377	AGCCCATATATGAGTTTCGCGTTACATAACTTACGGTAAATGGCCGCGCTGACTGACCG	436
DB	377		
QY	541	CCCAACGACCCCGCCCATATGAGTCAATTAAGCACTATGTTCCCATAGTACGCCAATA	600
DB	541		
QY	437	CCCAACGACCCCGCCCATATGAGTCAATTAAGCACTATGTTCCCATAGTACGCCAATA	496
DB	437		
QY	601	GGGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACATTTGGCAGTA	660
DB	601		
QY	497	GGGACTTTCCATTTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCACATTTGGCAGTA	556
DB	497		
QY	661	CATCAAGTGTATCATATGCGAAGTCCGCCCTTATTGAGCGTCAATGACGGTAAATGGCCC	720
DB	661		
QY	557	CATCAAGTGTATCATATGCGAAGTCCGCCCTTATTGAGCGTCAATGACGGTAAATGGCCC	616
DB	557		
QY	721	GCGTGCATTTATGCCAGTACATGACCTTACGGGACTTTCCCTACTTTGGCAGTACATCTAC	780
DB	721		
QY	617	GCGTGCATTTATGCCAGTACATGACCTTATGGGACTTTCCCTACTTTGGCAGTACATCTAC	676
DB	617		
QY	781	GTATTAGTCACTCGCTATTACCATGGTGTGATCGGTTTTTGGCAGTACACCAATGGCGGTGA	840
DB	781		
QY	677	GTATTAGTCACTCGCTATTACCATGGTGTGATCGGTTTTTGGCAGTACATCAATGGCGGTGA	736
DB	677		
QY	841	TAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTCCACCCCATGACGTCAATGGGAGTTTG	900
DB	841		
QY	737	TAGCGGTTTGACTCAGCGGGATTTCCAAAGTCTCCACCCCATGACGTCAATGGGAGTTTG	796
DB	737		
QY	901	TTTTGGCACCACCAATCAACGGGACTTTCCAAATGTGCTAATAACCCCGCCCGTTGACG	960
DB	901		
QY	797	TTTTGGCACCACCAATCAACGGGACTTTCCAAATGTGCTAATAACCCCGCCCGTTGACG	856
DB	797		
QY	961	CAAAATGGCGGTAGGGGTGTACCGTGGGAGGTCTATATAGCAGAGCTCGTTTGTGTAAC	1020
DB	961		
QY	857	CAAAATGGCGGTAGGGGTGTACCGTGGGAGGTCTATATAGCAGAGCTCGTTTGTGTAAC	916
DB	857		
QY	1021	CGTCAGATCCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAGACACCGGGAC	1080
DB	1021		
QY	917	CGTCAGATCCGCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAGACACCGGGAC	976
DB	917		
QY	1081	CGATCCAGCTCCGCGCGGGACCGGTGATTTGGAACCGGATTCCTCCGTCGCAAGAGT	1140
DB	1081		
QY	977	CGATCCAGCTCCGCGCGGGACCGGTGATTTGGAACCGGATTCCTCCGTCGCAAGAGT	1036
DB	977		
QY	1141	GAGCTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1200
DB	1141		
QY	1037	GAGCTAAGTACCGCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1096
DB	1037		

QY	1201	CTGTTTTTGGCTTGGGGCTATACACCCCGC--TCCTTATGCTATAGGTGATGGTATAGC	1259
DB	1201		
QY	1097	CTGTTTTTGGCTTGGGGCTATACACCCCGCTCTCTTATGCTATAGGTGATGGTATAGC	1156
DB	1097		
QY	1260	TTAGCCTATAGGTGTGGGTATTGTGACCACTATTGTGACCACTCCCTTATTGGTGACGATCT	1319
DB	1260		
QY	1157	TTAGCCTATAGGTGTGGGTATTGTGACCACTATTGTGACCACTCCCTTATTGGTGACGATCT	1216
DB	1157		
QY	1320	TTCCATTAATAATCAATAACATGGCTCTTTGGCCAACTATCTCTATTGGCTATATGCCA	1379
DB	1320		
QY	1217	TTCCATTAATAATCAATAACATGGCTCTTTGGCCAACTATCTCTATTGGCTATATGCCA	1276
DB	1217		
QY	1380	ATACTCTGCTCTCAGAGCTCAGACGACTCTGTATTTTACAGATGGGT--CCATTT	1438
DB	1380		
QY	1277	ATACTCTGCTCTCAGAGCTCAGACGACTCTGTATTTTACAGATGGGT--CCATTT	1336
DB	1277		
QY	1439	ATTATTTACAAATTCACATATACAAACCGCGTCCCCCGTCCCGCAGTTTTTATATAA	1498
DB	1439		
QY	1337	ATTATTTACAAATTCACATATACAAACCGCGTCCCCCGTCCCGCAGTTTTTATATAA	1396
DB	1337		
QY	1499	CATAGCGTGGGATCTCC--GACATCTCGGGTACGTTTCCGGACATGGGCTCTTCTCCG	1555
DB	1499		
QY	1397	CATAGCGTGGGATCTCCACGCGAACTCTCGGGTACGTTTCCGGACATGGGCTCTTCTCCG	1456
DB	1397		
QY	1556	GTAGCGGCGGAGCTTCCACATCCGAGCCCTGCTCCATCCGTCAGCGGCTCATGGTCG	1615
DB	1556		
QY	1457	GTAGCGGCGGAGCTTCCACATCCGAGCCCTGCTCCATGCTCCAGCGGCTCATGGTCG	1516
DB	1457		
QY	1616	TCGGCAGCTCTTGTCTTAAACAGTGGAGCCAGACTTTAGGCACAGCACAATGCCACCA	1675
DB	1616		
QY	1517	TCGGCAGCTCTTGTCTTAAACAGTGGAGCCAGACTTTAGGCACAGCACAATGCCACCA	1576
DB	1517		
QY	1676	CCACAGTGTGCCGACAAAGCCCGTGGGGTATGTCTGAAATAGCTCGGAG	1735
DB	1676		
QY	1577	CCACAGTGTGCCGACAAAGCCCGTGGGGTATGTCTGAAATAGCTCGGAG	1636
DB	1577		
QY	1736	ATTGGGCTCGACCTG--GACGAGATGGAGACTTTAAGCAGCGGCAGAGAGATGCAG	1794
DB	1736		
QY	1637	ATTGGGCTCGACCGCTGACGAGATGGAGACTTTAAGCAGCGGCAGAGAGATGCAG	1696
DB	1637		
QY	1795	GCAGCTGAGTGTGTGTATTCTGATAAGAGTCAAGGTAACTCCCGTTGCGGTGCTTTAA	1854
DB	1795		
QY	1697	GCAGCTGAGTGTGTGTATTCTGATAAGAGTCAAGGTAACTCCCGTTGCGGTGCTTTAA	1756
DB	1697		
QY	1855	CGGTGAGGCGCAGTGTAGTCTCAGCAGTACTCGTTGCTGCCGCGCGCCACACAGACATA	1914
DB	1855		
QY	1757	CGGTGAGGCGCAGTGTAGTCTCAGCAGTACTCGTTGCTGCCGCGCGCCACACAGACATA	1816
DB	1757		
QY	1915	ATAGCTGACAGACTAACAGACTGTTCTTCCATGGGTCTTTTCTGCACTACCGTCGTC	1974
DB	1915		
QY	1817	ATAGCTGACAGACTAACAGACTGTTCTTCCATGGGTCTTTTCTGCACTACCGTCGTC	1876
DB	1817		
QY	1975	GACGAATTCAGCA--ATCATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGC	2031
DB	1975		
QY	1877	GACAGTGTATCAGATATCGCGCGCTCTAGACACCATGCGCGTGAAGGAGAGTACC	1936
DB	1877		
QY	2032	TGTGTGAGCAGTCTTGTGTTTCCGCGAGCGCTAGCGGAAACCCACGTCACCGGGGGAAGTG	2091
DB	2032		
QY	1937	AGCACCTGTGGCGCTGGGGCTGGCGCTGGGGCACCATGCTGTGGGCATGCTGATGATCT	1996
DB	1937		
QY	2092	CGGCGCACACTGTGTGGATTGTGTAGCTCTCTCGCACAGGCGCCACAGAGAGCTCC	2151
DB	2092		
QY	1997	GACGCGCCACCGAGGCG-----TAGCAGCTGAAAGTGCACCGACCTGAAGAACACA	2047
DB	1997		
QY	2152	AGCTGATCAACACCAACCGCAGTTGGCACCTCAATAGCACCGGCCCTTGAACTGCAATGATA	2211
DB	2152		
QY	2048	CCAAACACCAACAGCAGCAGCGCGCGCATGATCATGGAGAAGGGCGAGATCAAGAACTGCA	2107
DB	2048		
QY	2212	GCTCAACACCGCGCTGTGGCAGGGCTTTTCTATACCAAGAGTTCAACTCTTCAAGCT	2271
DB	2212		
QY	2108	GCTTCAACATCAGCAGCAGCATCCGCGGCAAGGTGCAAGAGGATACGCTTCTTCTACA	2167
DB	2108		
QY	2272	GTCTGAGAGGCTAGCCAGCTCCGACCCCTTACCGATTTTTGACAGGCGCTGGGGCCCTA	2331
DB	2272		

Db 2168 AGTGGACATCATCCCATCGAACACACACACACAGCTACAGCCTGACAGCTGCAACA 2227
Qy 2332 TCAGTTATGCAACGGAAGCGGCCCGACAGCGCCCTACTGCTGGCACTACCCCCA - 2390
Db 2228 CCAGCGTATACCCAGGCTCTCCCAAGGTGAGCTTCGAGCCATCCCCAACCACTACT 2287
Qy 2391 --AAACCTTCGGTATTGTGCCCGCGAAGAGTGTGTGTGTCGGGTATATTGTCTCACTC 2448
Db 2288 GCGCCCCCGCGCTTCGCCATCTGAAGTGAAGGACAAAGTTCAACGGCAAGGGCC 2347
Qy 2449 CAGCCCCCGTGGT-----GTGGGAACGACGAGCTGGCGGGCGGCCCACTCAGCTGG 2504
Db 2348 CTTGCACCAACGAGCAGCTGAGTGCACCCACGCGCATCCGCGCGTGGTGAACCC 2407
Qy 2505 GGTGAAATGATACGAGCTCTTCGCTTAAATAACACAGGCGCCGCTGGGCAATTGG 2564
Db 2408 AGTGTGCTGAACGGTAGCTGGCCGAGGAGAGGTGTGTATCCGAGCGCTAACTTCG 2467
Qy 2565 TTCGGTTGTACCTGGATGAACCTGAATTCACCAAGTGTGCGGAGCGCCT--CCITT 2622
Db 2468 CCGACACGCCAAGGTATCATCTGTCAGCTGAACGAGAGGTGGAGATCAATGCACCC 2527
Qy 2623 GTGTATCGAGGGGGGGGCAACAAACCTGCACTGCCCACTGATTTGTTCCGCAAGC 2682
Db 2528 GCGCCCAACAAACACCGCAAGAGCATCCACATCGCGCGCGCGCTTCTACACCA 2587
Qy 2683 ATCCGAGGCGCATATCTCGGTGGGCTCGGTCCTGGATCACACCCAGTGGCTGG 2742
Db 2588 CCGCGAGATCATCGGCGCATCGCCAGCGCCCACTGCAACCTGAGCGCGCCCAAGTGA 2647
Qy 2743 TCAGTACCCGATAGGCTTTGGCATTTATCTGTACCATCAACTACACCATATTTAAAA 2802
Db 2648 ACACACCTTGACAAAGATCGTATCAAGCTGCGGAGAGTTGCGCAACAAACCATCG 2707
Qy 2803 TCAGGATGTAAGTGGGAGGGTGAACACAGGTGGAGTGGCTGCAACTGGAACGGGG 2862
Db 2708 TGTTCAAGCACAGCGCGGGGACCCCGAGATCGTGACCCACAGCTTCAACGTGGCGG 2767
Qy 2863 GCGAAGTTCGGATCTGGAAGATAGGACAGGTCCGAGATCGATATGGAACATCACAT 2922
Db 2768 GCGAGTTCTTCTAGTGAACAGACCCAGCTGTTTCAACAGCACCTGTGTTCAACAGCACCT 2827
Qy 2923 CAGGATTCCTAGGACCCCTGCTGTGTACAGCGGGGTTTTTCTTTGACAAAGATCC 2982
Db 2828 GGAGCACCGAGGCGCAACAAC-----ACCGAGGCGACGGACACCATCACCTGCCCT 2881
Qy 2983 TCACAATACCGCAGAGTCTAGACTCGGTGGATCTCTCAATTTTCTAGGGGATCTC 3042
Db 2882 GCGCATCAAGCAGATCATCAACATGTGGCAGAAGGTGGGCAAGGCCATGTACGCCCCC 2941
Qy 3043 CCTGTGTCTTGGCCAAATTCGAGTCCCAAGTCCCACTCAATCACTCAACACCTCTGTC 3102
Db 2942 CCATCAGCGCCAGATCCGCTGCAGCAGCAACATCAACCGCCCTGCTGCTGACCCGAGC 3001
Qy 3103 CTCCAATTTGCTGTTATCGTGTGATGTGTGCGGCGTTTTATCATATTCTCTTCA 3162
Db 3002 GCGCACACACAAACGAGCGAGATCTTCGTCGCGGGGGGACATCGCGGACA 3061
Qy 3163 TCCTGTGCTATGCTCATCTTATTATGTTTCTTGTGATATC-----AAGTATGTTG 3218
Db 3062 ACTGGCGCAGCAGCTGTCAAGTGTCAAGGTGTGAAGATCGAGCCCTGGCGGTGGGCC 3121
Qy 3219 CCGTTTGTCTTATTCAGGATCAACAAACAGTACGGGACCATCAAAACCTGTC 3278
Db 3122 CCACCAAGGCCAAGCTTACCGTCCAGGCCCGCAGCTGTGTGAGCGGCATCTGCAGCAGC 3181
Qy 3279 ACAGTCTCTCAAGCAACTCTATGTTTCCCTCATGTTGCTGTACAAACCTACGGAT 3338
Db 3182 AGAACAACTGTGCGGCCCATCGAGGCCAGGACACCTGTCTGACGTACCGTGTGGG 3241
Qy 3339 GGAATTTGCACTGTATTTCCATCCCATCGTCTCTGG-----GTTTTGCAAAATACCTATGG 3395

Db 3242 GCATCAAGCAGCTGCGAGCGCCGCAACCTGGCGTGGAGCGCTACCTGAAGACAGCAGC 3301
Qy 3396 GAGTGGGCGCTCAGTCCGTTTCTTGGCTCAGTTTACTAGTGCAATTTGTTAGTGGTTC 3455
Db 3302 AGCTCAGCAGATCTGGAACCAACACCTGGATGAGTGGGACCGGAGATCAACAACT 3361
Qy 3456 GTAGGGCTTTCCCCCACTGTTTGGCTTTTCTAGCTATATGGATGATGTTGTTGGGCCA 3515
Db 3362 ACACGAGCTGATCCACAGCTGATCGAGGAGAGCCAGAACACGACGAGAGAAGCAGC 3421
Qy 3516 AGTCTGTACAGCATCTGAGTCCCTTTATACCGCTGTTACCAATTTTCTTTGTTCTCTGG 3575
Db 3422 AGGAGCTGTGGAGCTGGACAAGTGGGCCAGCTGTGGAACTGGTTCAACATCACCAACT 3481
Qy 3576 GTATACATTTAAGAAATTCAGACTCGAGCAAGTCTAGAAAGGCGCGCAAGATATCAAGGA 3635
Db 3482 GGTGTGTATACATCAAGCTGTTTCTATCATGATCGTGGCGGCTGTGGGCTCGGCATCG 3541
Qy 3636 TCCACTACGCG--TTAGAGCTCGCTGATCAGCTCGACTGCTGCTTCTAGTTGCAGCCAT 3694
Db 3542 TGTTCCGCGTGTGAGCATCTGAGGATCCAGATCTGCTGTGCTTCTAGTTGCAGCCAT 3601
Qy 3695 CTGTTGTTGCCCTCCCGCTGCCCTTCTTGAACCTGGAAGTGCACCTCCACCTGCTCC 3754
Db 3602 CTGTTGTTGCCCTCCCGCTGCCCTTCTTGAACCTGGAAGTGCACCTCCACCTGCTCC 3661
Qy 3755 TTTCTTAATAAATAGAGAAATTCATCGCATCTGCTGAGTAGTGTGTCATTCTATTCTGG 3814
Db 3662 TTTCTTAATAAATAGAGAAATTCATCGCATCTGCTGAGTAGTGTGTCATTCTATTCTGG 3721
Qy 3815 GGGTGGGTGGGCGAGGACAGCAAGGGGAGGATTTGGGAAGACAATAGCAGGCTGCTG 3874
Db 3722 GGGTGGGTGGGCGAGGACAGCAAGGGGAGGATTTGGGAAGACAATAGCAGGCTGCTG 3781
Qy 3875 GGAGC----- 3880
Db 3782 GGGATCGGTGGGCTCTATGGGTACCCAGTGTGAAGAATTGACCCGGTTCTCTCTGGG 3841
Qy 3881 ----- 3880
Db 3842 CCAAGAAAGACGAGGACATCCCTTCTCTGTGACACACCTGTCCACGCCCTGTTCT 3901
Qy 3881 ----- 3880
Db 3902 TAGTTCCAGCCCCACTCATAGGACACTCATAGCTCAGGAGGCTCCGCTTCAATCCCAC 3961
Qy 3881 ----- 3880
Db 3962 CCGCTAAAGTACTTTGGAGGGTCTCTCTCCCTCATCAGCCCAACCAACCTAGC 4021
Qy 3881 ----- 3880
Db 4022 CTCCAAGGTGGGAAGAAATTAAAGCAAGTAGGCTATTAAAGTGCAGAGGAGAGAAAT 4081
Qy 3881 ----- 3880
Db 4082 GCCTCCAAACATGTGAGGAAGTAATGAGAAATCATAGAATTTTAAGGCCATCATGSCCT 4141
Qy 3881 ----TCTTCGCTTCTCGCTCACTGACTCGTGGCTCGGTCTGGTCTGGGCGGAGCG 3937
Db 4142 TAATCTCGGCTTCTCGCTCACTGACTCGCTCGCTCGGTCTGGTCTGGGCTGGGCGG 4201
Qy 3938 GTATCAGCTCACTCAAGGGGTAAATACGGTTATCCACAGAATCAGGGATTAACGAGGA 3997
Db 4202 GTATCAGCTCACTCAAGGGGTAAATACGGTTATCCACAGAATCAGGGATTAACGAGGA 4261
Qy 3998 AAGAAATGTGAAGAAAGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTGTGCTG 4057
Db 4262 AAGAAATGTGAAGAAAGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTGTGCTG 4321
Qy 4058 GCGTTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAG 4117
Db 4322 GCGTTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAG 4381

4118 AGGTGGCGAAACCCGACAGAGCTATATAAGATACCAGGCGTTTCCCGCTCGAAGCTCCCTC 4177
Db |||||||
4382 AGGTGGCGAAACCCGACAGAGCTATATAAGATACCAGGCGTTTCCCGCTCGAAGCTCCCTC 4441
Qy |||||||
Db |||||||
4178 GTGCGCTCTCCGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCG 4237
Db |||||||
4442 GTGCGCTCTCCGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCG 4501
Qy |||||||
Db |||||||
4238 GGAAGCGTGGCGCTTTCTCAATGCTCACGCTGATGATCTCAGTTTCGGTGTAGTGCCTT 4297
Db |||||||
4502 GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGTGCCTT 4561
Qy |||||||
Db |||||||
4298 CGTCTCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCC 4357
Db |||||||
4562 CGTCTCAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCC 4621
Qy |||||||
Db |||||||
4358 GGTAACTATCGTCTGAGTCCAAACCCGTTAAGACAGACTTATCGCCACTGGCAGCAGCC 4417
Db |||||||
4622 GGTAACTATCGTCTGAGTCCAAACCCGTTAAGACAGACTTATCGCCACTGGCAGCAGCC 4681
Qy |||||||
Db |||||||
4418 ACTGTAACAGGATTAAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGG 4477
Db |||||||
4682 ACTGTAACAGGATTAAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGG 4741
Qy |||||||
Db |||||||
4478 TGGCCTAACTACGGCTACACTAGAAGACAGATATTTGGTATCTGCGCTCTGCTGAAGCCA 4537
Db |||||||
4742 TGGCCTAACTACGGCTACACTAGAAGACAGATATTTGGTATCTGCGCTCTGCTGAAGCCA 4801
Qy |||||||
Db |||||||
4538 GTTACCTTCGGAAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCGCTGTAGC 4597
Db |||||||
4802 GTTACCTTCGGAAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCGCTGTAGC 4861
Qy |||||||
Db |||||||
4598 GGTGGTTTTTTTGTGTAAGCAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAAGAT 4657
Db |||||||
4862 GGTGGTTTTTTTGTGTAAGCAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAAGAT 4921
Qy |||||||
Db |||||||
4658 CCTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGGATT 4717
Db |||||||
4922 CCTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAAGAAACTCAGTTAAGGGATT 4981
Qy |||||||
Db |||||||
4718 TTGCTCATGAGATTATCAAAAAGATCTTCACCTAGATCTTTTAAATTAATAATGAAGT 4777
Db |||||||
4982 TTGCTCATGAGATTATCAAAAAGATCTTCACCTAGATCTTTTAAATTAATAATGAAGT 5041
Qy |||||||
Db |||||||
4778 TTTAAATCAATCTAAAGTATATATGATGAATCTTGTCTGACAGTTTACCAATGCTTAATC 4837
Db |||||||
5042 TTTAAATCAATCTAAAGTATATATGATGAATCTTGTCTGACAGTTTACCAATGCTTAATC 5101
Qy |||||||
Db |||||||
4838 AGTGAGCAGCTATCTCAGGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTC 4894
Db |||||||
5102 AGTGAGCAGCTATCTCAGGATCTGTCTATTTTCGTTTCATCCATAGTTGCTGACTC 5158

RESULT 11

US-10-359-120-15
; Sequence 15, Application US/10359120
; Publication No. US20040033487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKRABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; TITLE OF INVENTION: ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 6505
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
US-10-359-120-15

Query Match 38.8%; Score 2283.4; DB 17; Length 6505;
Best Local Similarity 69.5%; Pred. No. 0;
Matches 3656; Conservative 0; Mismatches 1061; Indels 546; Gaps 19;
Qy 1 TCGCGCGTTTCGGTGATGACCGTGAAACCTCTGACACATGCGAGCTCCCGGAGACGGTCA 60
Db 1 TCGCGCGTTTCGGTGATGACCGTGAAACCTCTGACACATGCGAGCTCCCGGAGACGGTCA 60
Qy 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGGTG 120
Db 61 CAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGGTG 120
Qy 121 TTGGCGGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
Db 121 TTGGCGGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
Qy 181 ACCATATGAAGCTTTTTCGAAAAGCCTAGGCGCTCCAAAAAGCCTCCTCACTACTTCTGG 240
Db 181 ACCATATGCGG-----TGTGAATACCGCACA 207
Qy 241 AATAGCTCAGAGCGCGAGCGGCTCGGCTCTGTCATTAATAAAAAAATTAGTCAGCCA 300
Db 208 GATGCGTAAGGAGAAAATACCGCATCAG----- 235
Qy 301 TGGGCGGAGATGGCGGAACTGGGCGGGAGGGAATTATTGGCTATTGGCCATTGCAAT 360
Db 236 -----ATTGGCTATTGGCCATTGCAAT 256
Qy 361 AGCTTGTATCTATATATAATAATGTACATTTATATTGGCTCATGTCCAATATGACGCCA 420
Db 257 AGCTTGTATCTATATAATAATGTACATTTATATTGGCTCATGTCCAATATGACGCCA 316
Qy 421 TGTGACATTTGATTTGCTAGTATTAAATAGTAACTCAATTAACGGGTCTAGTTTCAAT 480
Db 317 TGTGACATTTGATTTGCTAGTATTAAATAGTAACTCAATTAACGGGTCTAGTTTCAAT 376
Qy 481 AGCCCATATATGAGTTCCGCGTTACATAAATTCAGTAAATGGCCGCTGGCTGACCG 540
Db 377 AGCCCATATATGAGTTCCGCGTTACATAAATTCAGTAAATGGCCGCTGGCTGACCG 436
Qy 541 CCCAACGACCCCGCCCAATTTGACATTAATGAGTATGTTCCCATAGTAAACGCCAATA 600
Db 437 CCCAACGACCCCGCCCAATTTGACATTAATGAGTATGTTCCCATAGTAAACGCCAATA 496
Qy 601 GGGACTTTCATTTGACGTCATCAATGGGTGGAGTATTTACGGTAAACTGCCCACTGGCAGTA 660
Db 497 GGGACTTTCATTTGACGTCATCAATGGGTGGAGTATTTACGGTAAACTGCCCACTGGCAGTA 556
Qy 661 CATCAAGTGTATCATATGCCAAGTCCGCCCTTATTGAGTCAATGACCGTAAATGGCC 720
Db 557 CATCAAGTGTATCATATGCCAAGTCCGCCCTTATTGAGTCAATGACCGTAAATGGCC 616
Qy 721 GCGTGGCATTTGCCAGTACATGACCTTTACGGGACTTTTCTACTTTGGCAGTACATCTAC 780
Db 617 GCGTGGCATTTGCCAGTACATGACCTTTTGGGACTTTTCTACTTTGGCAGTACATCTAC 676
Qy 781 GTATTAGTCAATCGCTATTACCAATGGGTGGAGTGGCGTTTGGCAGTACCAATGGCGTGA 840
Db 677 GTATTAGTCAATCGCTATTACCAATGGGTGGAGTGGCGTTTGGCAGTACCAATGGCGTGA 736
Qy 841 TAGCGGTTTGACTCAGCGGGGATTTTCCAGTCTCCACCCCAATTTGACGTCATCAATGGGAGTTG 900

Db 737 ||||| TAGCGGTTGACTCAGCGGATTTCCAAAGTCTCCACCCATTGACGTCATGCGAGTTTG 796
Qy 901 ||||| TTTTGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 960
Db 797 ||||| TTTTGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCGTTGACG 856
Qy 961 ||||| CAAATGGCGGTAGGGTGTGACGTTGGGAGTCTATATAAGCAGAGTCTGTTAGTGAAC 1020
Db 857 ||||| CAAATGGCGGTAGGCGTGTACGTTGGGAGTCTATATAAGCAGAGTCTGTTAGTGAAC 916
Qy 1021 ||||| CGTCAGATCGCCTGGAGAGCCCAATCAACGCTGTTTGAACCTCATAGAGACACCGGGAC 1080
Db 917 ||||| CGTCAGATCGCCTGGAGAGCCCAATCAACGCTGTTTGAACCTCATAGAGACACCGGGAC 976
Qy 1081 ||||| CGATCAGCCTCCCGCGCGGGAACGTCATTTGGAACCGGATTCCTCGTGCACAGAGT 1140
Db 977 ||||| CGATCAGCCTCCCGCGCGGGAACGTCATTTGGAACCGGATTCCTCGTGCACAGAGT 1036
Qy 1141 ||||| GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1200
Db 1037 ||||| GACGTAAGTACCGCCTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA 1096
Qy 1201 ||||| CTGTTTTGGCTTGGGCGCTATACACCCCGC--TCCTTATGCTATAGGTGATGGTATAGC 1259
Db 1097 ||||| CTGTTTTGGCTTGGGCGCTATACACCCCGCTTCTTATGCTATAGGTGATGGTATAGC 1156
Qy 1260 ||||| TTAGCCTATAGGTGCGGTATTGACCAATATTGACCACTCCCTTATGGCTATATGCAATCT 1319
Db 1157 ||||| TTAGCCTATAGGTGCGGTATTGACCAATATTGACCACTCCCTTATGGCTATATGCAATCT 1216
Qy 1320 ||||| TTCATTACTAATPCCATAACATGGCTCTTTGCCACAACTATCTATTTGGCTATATGCCA 1379
Db 1217 ||||| TTCATTACTAATPCCATAACATGGCTCTTTGCCACAACTATCTATTTGGCTATATGCCA 1276
Qy 1380 ||||| ATACTCTGCTTTCAGAGACTGACAGGACTCTGTATTTTTCAGAGATGGGT--CAATTT 1438
Db 1277 ||||| ATACTCTGCTTTCAGAGACTGACAGGACTCTGTATTTTTCAGAGATGGGTCCCATTT 1336
Qy 1439 ||||| ATTATTTACAAATTCACATATACAAACGCGCTCCCGTCCCGTCCCGAGTTTATTTAA 1498
Db 1337 ||||| ATTATTTACAAATTCACATATACAAACGCGCTCCCGTCCCGTCCCGAGTTTATTTAA 1396
Qy 1499 ||||| CATAGCGTGGGATCTCC--GACATCTCGGTACGTGTTCCGGACATGGCTCTTCTCG 1555
Db 1397 ||||| CATAGCGTGGGATCTCCACGCGAATCTCGGTACGTGTTCCGGACATGGCTCTTCTCG 1456
Qy 1556 ||||| GTAGCGCGGAGCTTCCACATCCGAGCCTGTGTCCTCATCGGTCCAGCGGCTCATGTGCG 1615
Db 1457 ||||| GTAGCGCGGAGCTTCCACATCCGAGCCTGTGTCCTCATCGGTCCAGCGGCTCATGTGCG 1516
Qy 1616 ||||| TCGGCAGCTCCTTGCTCCTAACAGTCGGAGCCAGACTTAGGCACAGCAATGCCACCA 1675
Db 1517 ||||| TCGGCAGCTCCTTGCTCCTAACAGTCGGAGCCAGACTTAGGCACAGCAATGCCACCA 1576
Qy 1676 ||||| CCACCAAGTGTCCGCAACAGCGCGTGGGTAGGTATGTCTGAAATAGCTCGAG 1735
Db 1577 ||||| CCACCAAGTGTCCGCAACAGCGCGTGGGTAGGTATGTCTGAAATAGCTCGAG 1636
Qy 1736 ||||| ATTGGGCTCGCAC--CTGGACGCAAGATGGAAGACTTAAGGAGCGCGCAGGAAGATGCGAG 1794
Db 1637 ||||| ATTGGGCTCGCACGCGCTGACGAGATGGAAGACTTAAGGAGCGCGCAGGAAGATGCGAG 1696
Qy 1795 ||||| GCAGCTGAGTGTGTGTTCTGATAGAGTCAGAGTAACTCCCGTGGGTGCTGTTAA 1854
Db 1697 ||||| GCAGCTGAGTGTGTGTTCTGATAGAGTCAGAGTAACTCCCGTGGGTGCTGTTAA 1756
Qy 1855 ||||| CGGTGAGGCGCAGTGTAGTCTGAGCAGTACTCGTGTGCTCGCGCGCGCCACAGACATA 1914
Db 1757 ||||| CGGTGAGGCGCAGTGTAGTCTGAGCAGTACTCGTGTGCTCGCGCGCGCCACAGACATA 1816
Qy 1915 ||||| ATAGCTGACAGACTAAACAGACTGTTCTCTTCCATGGGTCTTTTCTCGAGTCAACCGTCGTC 1974

Db 1817 ||||| ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGTCTTTTCTGCACTCACCCTGTC 1876
Qy 1975 ||||| GACGAATTCAGACA--ATCATGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTC 2031
Db 1877 ||||| GACAGTGTGATCAGATATCGCGCGCTCTTAGACACCATGCGGTGAGGAGAGTACC 1936
Qy 2032 ||||| TGTGTGAGACAGTCTTCGTTTTCGCCAGCGCTAGCGAAAACCCACGTCACCGGGGAGTG 2091
Db 1937 ||||| AGCACCTGTGGCGCTGGGCTGCGCTGCGGCACCATGCTGCTGGGATGCTGATGATCT 1996
Qy 2092 ||||| CCGGCACACTGTGTCTGATTTTGTAGCTCTCTCCACAGCGCGCAGCAGCAACGTC 2151
Db 1997 ||||| GCAGCGCCACCGAG-----GCTAGCAGCTGAAGTGCACCGACCTGAAGACGACA 2047
Qy 2152 ||||| AGCTGATCAACACCAACGCGAGTTGCACCTCAATAGCAGCGGCTGAACTGCAATGATA 2211
Db 2048 ||||| CCAACACCAACAGCAGCAGCGCGCATGATCATGAGAGAGGCGAGATCAAGACTGCA 2107
Qy 2212 ||||| GCTCAACACCGGCTGTTTGGCAGGGCTTTTCTATCACCACAAAGTTCAACTCTTCAGGCT 2271
Db 2108 ||||| GCTTCAACATCAGCAGCAGCATCCGCGCAAGGTGCAGAAAGGAGTACGCTTCTTCTACA 2167
Qy 2272 ||||| GTCCTGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTTTGAACAGGGTGGGGCCTA 2331
Db 2168 ||||| AGCTGACATCATCCCATCGAACACGACACCAACAGCTACAGCTGACCGCTGCAACA 2227
Qy 2332 ||||| TCAGTTTATGCCAAGGAGCGCGCCGACGAGCGCCCTACTGCTGGCACTACCCCCAA 2391
Db 2228 ||||| CCAGCGTATACCCAGGCTGCGCCCAAGGTGAGCTTCCAGGCCATCCCAACACTACT 2287
Qy 2392 ||||| AACCTTTCGGTATTTGTCGCCGGAAGAGTGTGTGTGGTCCGGTATATTGCTTCACTCCCA 2451
Db 2288 ||||| GCGCCCCCGCGCTTCGCCCATCTCTGAA-----GTGCAAGGACAAGAATTCA 2335
Qy 2452 ||||| GCGCGTGTGTGGGAACGACGACAGTGGCGCGCGCCACCTACAGCTGGGGTGAAA 2511
Db 2336 ||||| ACGGCAAGGGCCCTGCAACAAAGTGCAGTGCACCCACCGCATCCGCCCCG 2395
Qy 2512 ||||| ATGATACGAGCTCTTCGTCCTTAAACAATACAGGCGCACCGCTGGGCAATTTGTTGCTT 2571
Db 2396 ||||| TGTGAGCACCCAGCTGCTGTGAAACGCTGAGCTGCGCGAGGAGGAGGTGATCG-- 2453
Qy 2572 ||||| GTACCTGGATGAATCAACTGGATTACCAAAAGTGTGCGAGCGCTCTTGTGTATCG 2631
Db 2454 ||||| ----CAGCGCTAATTCGCGCAACGCGCAAGGTGATCATCTGTGACGTGAACGAGA-- 2506
Qy 2632 ||||| GAGGGCGGGGAAACACACCTGTGACTGCCCACTGATTTGCTTCCGCAAGCATCCGAGC 2691
Db 2507 ||||| GCGTGGAGATCAACTGCAACCGCCCCCAACAACAACCGCGCAAGAGCATCCACATCGGC 2566
Qy 2692 ||||| CCACATCTCTCGGTGCGGCTCCGCTCCCTGGATCACACCCAGGTGCTGGTGGCTTACC 2751
Db 2567 ||||| CCGCGCGCTTCTTACACACCGGAGATCATCGCGACATCCCGCAGGCCACTTGCA 2626
Qy 2752 ||||| CGTATAGGCTTTGGCATTTATCTTGTACCATCAACTACACCATATTTAAATCAGGATGT 2811
Db 2627 ||||| AACTGAGCGCGCAAGTGAACGACACCTGAAACAAGATCGTGATCAAGCTCGCGAGC 2686
Qy 2812 ||||| AGTGGAGGGGTGCAACACAGCTGGAAGCTGCTGCAACTGGAACGCGGGGGAAGTT 2871
Db 2687 ||||| AGTTCCGCAA----CAAGACCATCTGTTTCAAGCACAGCAGCGCGCGCCCGAGATC 2742
Qy 2872 ||||| GCGATCTGGAAGTAGGAGAGTCCGAGATCGATATGAGAACATCATCATAGGATTC 2931
Db 2743 ||||| GTGACCAAGCTTCAACTGCGCGCGGAGTTCTTCTACTGCAACAGCAGCCAGCTGTTT 2802
Qy 2932 ||||| TAGGACCCCTGCTGCTTGTACAGGGGGGTTTTTCTTGTGTGACAAAGATCTTCAATAC 2991
Db 2803 ||||| AACAGCACTGTTCAACAGCACTTGGACACCGAGGGCAGCAACAACCGAGGGCAGC 2862
Qy 2992 ||||| CGCAGAGTCTAGACTGCTGCTGAGCTTCTCTCAATTTTCTAGGGGATCTTCCGCTGTC 3051
Db 2863 ||||| GACACCATCACCTGCGCTGCGCATCAAGCAGATCATCAACATGTGGCAGAGGTGGC 2922

QY	3052	TTGGCCAAAATTGCGAGTCCCCAACTCCAAATCACTACAAACCTCCTCTCTCCCAATTT	3111	QY	3879	-----	3878
Db	2923	AAAG-----CCATGTACGCCCCCCCATCAGCGGCAGATCC	2959	Db	3964	AGAAATTAAAGCAAGATAGGCTATTAAGTGCAGAGGAGAGAAATGCTCCAAATGTG	4023
QY	3112	GTCCCTGGTTATCGCTGGATGTGTCGGGGTTTTATCATATATTCCTCTTCATCCCTGCTGC	3171	QY	3879	-----GCTCTTCGGCTTC	3891
Db	2960	GCTGCAGCAGCAACATCACCGGCTCTGCTGACCCGCGACGGCGCAACAGCAACAACG	3019	Db	4024	AGGAAGTAATGAGAGAAATCATAGAAATTTAAGGCCATCATGGCCTTAATCTTCGGCTTC	4083
QY	3172	TATGCTCATCTCTTTATTTATTTGTTCTTCTGGAATTAACAAGGTATGTTC--CGTGTGTCTCT	3230	QY	3892	CTCGCTCACTGACTCGCTCGCTCGGTCTCGGTCTGGGCGAGCGGTATCAGCTCACTC	3951
Db	3020	AGAGCGAGATCTTCCTCGCGGGCGCGGACATGCGGACAACTGGCGCAGGAGCTGT	3079	Db	4084	CTCGCTCACTGACTCGCTCGCTCGGTCTCGGTCTGGGCGAGCGGTATCAGCTCACTC	4143
QY	3231	CTAATTTCCAGGATCAACAACAACAGTACGGGACCAATGCAAACTTCGACGACTCTCTGT	3290	QY	3952	AAAGGGGTAAATACCGTTATCCACAGAAATCAGGGGATACGAGAGAAAGAAATGTGAGC	4011
Db	3080	ACAAGTACAAGGTGTGAAGATCGAGCC-----CTGGGCGTGGCCCAACCAAGG	3130	Db	4144	AAAGCGGTAAATACCGTTATCCACAGAAATCAGGGGATTAACGAGGAAACATGTGAGC	4203
QY	3291	CAAGGCAACTATGTTTCCCTCATGTTCTGTACAAACCTACGGATGGAATTCACCC	3350	QY	4012	AAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGCCCGGTTGCTGGCGTTTTTCCATAG	4071
Db	3131	CAAAGCTTACCGTTCAGGCCCGCGACGTCTGAGCGGCATCGTGCAGCAGCAACAAC	3190	Db	4204	AAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGCCCGGTTGCTGGCGTTTTTCCATAG	4263
QY	3351	TGTATTTCCCATCCCATCGTCTGGGCTTTTCGCAAAATACCTATGGGAGTGGGCTCAGTC	3410	QY	4072	GCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCAAAACC	4131
Db	3191	TGCTCGCGCCATCGAGGCCCGACGACCACTGCTGCAGCTGACCGTGTGGGCAATCAAGC	3250	Db	4264	GCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCAAAACC	4323
QY	3411	CGTTTCTCTTGGCTCAGTTTACTAGTGGCAATTTGTTCACTGGTTCTGATGGGCTTTCCCCC	3470	QY	4132	GACAGGACTATAAGATACCAAGCGGTTTTCCCTCGAAAGCTCCCTCGTGGCTCTCCTGT	4191
Db	3251	AGCTGCAGGCC--CGCAACCTGGCCGTGGAGCGGTACTCTGAAGACCAAGCAGCTCGAG	3309	Db	4324	GACAGGACTATAAGATACCAAGCGGTTTTCCCTCGAAAGCTCCCTCGTGGCTCTCCTGT	4383
QY	3471	ACTGTTTGGCTTTCAGCTATATGGATGATGTGTATTTGGGGCCAGTCTGTACAGCATC	3530	QY	4192	TCCGACCTTCCGCTTACCGGATACCTGTCCGCCCTTTCTCCCTTGGGAAGCGTGGCGCT	4251
Db	3310	CAGATCTGGAACCAACACCTGGATGGAGTGGGACCGCGAGATCAACAATACACCAAC	3369	Db	4384	TCCGACCTTCCGCTTACCGGATACCTGTCCGCCCTTTCTCCCTTGGGAAGCGTGGCGCT	4443
QY	3531	GTAGTTCCTTTATACCGCTGTGTACCAATTTTCTTTTGTCTCTGGGTATACATTTAAGAA	3590	QY	4252	TTCTCAATGCTCACTGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTCCAAGTGGG	4311
Db	3370	CTGATCCAC-----AGCCTGATCGAGAGAGCCAGAACACGACGAGAAACAGCAG	3423	Db	4444	TTCTCATAGTCACTGCTGTAGGTATCTCAGTTTCGGGTAGGTGCTTCGCTCCAAGTGGG	4503
QY	3591	TTCAGACTCGAGCAAGCTAGAAAGCGCGCCAGATATCAAGGATCCACTAGCGGTTAG	3650	QY	4312	CTGTGTGCAGCAACCCCGGTTACGCCGACCGCTCGCGCTTATCCGTTAACTATCGTCT	4371
Db	3424	GAGCTCTGGAGCTGACAAAGTGGGCGAGCTGTGGAATCTGTTCAACATCAACCACTGG	3483	Db	4504	CTGTGTGCAGCAACCCCGGTTACGCCGACCGCTCGCGCTTATCCGTTAACTATCGTCT	4563
QY	3651	AGCT--CGCTGATCAGCTCGACTGCTCTAGTTGCGAGCACTCTGTGTTGSCCCC	3708	QY	4372	TGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGAT	4431
Db	3484	CTGTGTGAGGATCCAGATCTGCTGTGCTCTTAGTTGCGAGCACTCTGTGTTGSCCCC	3543	Db	4564	TGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGAT	4623
QY	3709	TCCCCCGTCTCTTCCCTGGAAGGTGGCACTCCCACTGTCTCTTCTCTTAATAAAAT	3768	QY	4432	TAGCAGAGCAGGATATGATGGCGGTCTACAGATTTCTTGAAGTGGTGCCCTAACTACGG	4491
Db	3544	TCCCCCGTCTCTTCCCTGGAAGGTGGCACTCCCACTGTCTCTTCTCTTAATAAAAT	3603	Db	4624	TAGCAGAGCAGGATATGATGGCGGTCTACAGATTTCTTGAAGTGGTGCCCTAACTACGG	4683
QY	3769	GAGGAAATTGCATCGCATTTGTCTGAGTAGGTGTCAATTTCTGCGGGGTGGGGTGGG	3828	QY	4492	CTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAA	4551
Db	3604	GAGGAAATTGCATCGCATTTGTCTGAGTAGGTGTCAATTTCTGCGGGGTGGGGTGGG	3663	Db	4684	CTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGAAA	4743
QY	3829	CAGGACAGAGGGGAGGATTTGGGAAGACAATAGCAGGCATGCTGGGATCGGGTGGG	3878	QY	4552	AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGTGTAGCGGTGGTTTTTGT	4611
Db	3664	CAGGACAGAGGGGAGGATTTGGGAAGACAATAGCAGGCATGCTGGGATCGGGTGGG	3723	Db	4744	AAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGTGTAGCGGTGGTTTTTGT	4803
QY	3879	-----	3878	QY	4612	TTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCTTTTC	4671
Db	3724	TCTATGGGTACCCAGGTGTGAAGATTTGACCCGGTTCTCTCTGGGCCAGAAAGACG	3783	Db	4804	TTGCAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCTCTTGAATCTTTTC	4863
QY	3879	-----	3878	QY	4672	TACGGGTCTGACGCTCAGTGGAAACCAACCTCAGCTTAAAGGATTTTTCGTCATGAGATT	4731
Db	3784	GCACATCCCTTCTCTGTGACACACCTGTTCACAGCCCTCGTCTTAGTTCCAGCCCCA	3843	Db	4864	TACGGGTCTGACGCTCAGTGGAAACCAACCTCAGCTTAAAGGATTTTTCGTCATGAGATT	4923
QY	3879	-----	3878	QY	4732	ATCAAAAAAGGATCTTTCACCTAGATCTTTTAAATTAATAAATGAAGTTTTTAAATCAATCTA	4791
Db	3844	CTCATAGGACATCATAGCTCAGGAGGGCTCGCCCTTCAATCCACCCCGTAAAGTACTTT	3903	Db	4924	ATCAAAAAAGGATCTTTCACCTAGATCTTTTAAATTAATAAATGAAGTTTTTAAATCAATCTA	4983
QY	3879	-----	3878	QY	4792	AAGTATATATAGTAAAATTGCTGTGACAGTTTACCAATGCTTAAATCAGTGGGACCTAT	4851
Db	3904	GGAGCGGTCTCTCCCTCCCTCATCAGCCCAACCAACCAACCTAGCCTCCAAAGAGTGGGA	3963	Db	4984	AAGTATATATAGTAAAATTGCTGTGACAGTTTACCAATGCTTAAATCAGTGGGACCTAT	5043
				QY	4852	CTCAGCGATCTGTCTATTATTTCGTTTCATCCATAGTTTCCTGACTC	4894


```
Db 5044 CTCAGCGATGCTGCTATTTGCTTCATCCATAGTGCCTGACTC 5086
|||||
RESULT 12
US-10-359-120-34
; Sequence 34, Application US/10359120
; Publication No. US20040033487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKRABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 6460
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
US-10-359-120-34

Query Match 38.8%; Score 2283; DB 17; Length 6460;
Best Local Similarity 69.9%; Pred. No. 0;
Matches 3595; Conservative 0; Mismatches 1195; Indels 355; Gaps 24;

Qy 1 TC CGCGGTTTCGCTGATGACGGTGAAACCTCTGACACATGACGTCCCGGAGACGGTCA 60
Db 1 TC CGCGGTTTCGCTGATGACGGTGAAACCTCTGACACATGACGTCCCGGAGACGGTCA 60

Qy 61 CAGCTTGCTCTGTAAGCGGATGCGGGAGACAGACAGCCCTCAGGGCGCGTCAAGCGGGTG 120
Db 61 CAGCTTGCTGTAAGCGGATGCGGGAGACAGACAGCCCTCAGGGCGCGTCAAGCGGGTG 120

Qy 121 TTGGCGGGTGTGGGGTGGCTTAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
Db 121 TTGGCGGGTGTGGGGTGGCTTAACTATCGGCATCAGAGCAGATTGTACTGAGAGTGC 180

Qy 181 ACCATATGAAGCTTTTGGCAAAAGCCTAGCGCTCCAAAAGCCTCCTCACTACTTCTGG 240
Db 181 ACCATATGGG-----TGTGAAATACCGCACA 207

Qy 241 AATAGCTCAGAGCGAGGGCGCTCGGCTCTGCTGATAATAAAAAAATTAGTCAGCCA 300
Db 208 GATGCGTAAGGAGAAATACCGCATCAG-----235

Qy 301 TGGGCGGAGAAATGGGCGGAACTGGGCGGGAGGAATTAATGGCTATTGGCCATTGCAAT 360
Db 236 -----ATTGGCTATTGGCCATTGCAAT 256

Qy 361 ACCTTGATCTATATCATATATGATACATTTATATTGGCTCATGTCCAAATATGACGCCA 420
Db 257 ACCTTGATCTATCATATATGATACATTTATATTGGCTCATGTCCAAATATGACGCCA 316

Qy 421 TGTTCATTTGATTTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTTCA 480
Db 317 TGTTCATTTGATTTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTTCA 376

Qy 481 AGCCCATATATGAGTTTCCCGTTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCG 540
Db |||||
```

1616 TCGGAGCTCCTTGCTCTTAACAGTGGAGGCGAGACTTAGGCGACAGCAACAATGCCACCA 1675
1517 TCGGAGCTCCTTGCTCTTAACAGTGGAGGCGAGACTTAGGCGACAGCAACAATGCCACCA 1576
1676 CCACCAAGTGTGCGCGCAAGAGGCGGTAGGGTATGTGCTGAAATAGCTCGGAG 1735
1577 CCACCAAGTGTGCGCGCAAGAGGCGGTAGGGTATGTGCTGAAATAGCTCGGAG 1636
1736 ATTGGGCTCGCAC - CTGGACGCGAGATGGAAGACTTAAGGCGAGCGGCGAGGAAGATGCGAG 1794
1637 ATTGGGCTCGCACGCGGTGACGCGAGATGGAAGACTTAAGGCGAGCGGCGAGGAAGATGCGAG 1696
1795 GCAGCTGAGTGTGTTGTTATCTGATTAAGAGTCAGAGGTAATCTCCGTTGGGTGCTGTTAA 1854
1697 GCAGCTGAGTGTGTTGTTATCTGATTAAGAGTCAGAGGTAATCTCCGTTGGGTGCTGTTAA 1756
1855 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTCGGCGCGCGCCACAGACATA 1914
1757 CGGTGAGGCGAGTGTAGTCTGAGCAGTACTCGTTGCTCGGCGCGCGCCACAGACATA 1816
1915 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGTCTTTTCTGCAGTCAACCGTCGC 1974
1817 ATAGCTGACAGACTAAACAGACTGTTCTTTTCCATGGGTCTTTTCTGCAGTCAACCGTCGC 1876
1975 GACGAATTCAGGAATCATGGATGC----- 1999
1877 GACAGTGTGATCAGATATACGGGCGGCTCTAGAATGCGCGTGAAGGAGAAGTACAGCA 1936
2000 ----- 1999
1937 CCTGTGGCGCTGGGCTGGCGTGGGGCACCATGCTGCTGGGCATGCTGATCATGTCGAG 1996
2000 -----AATGAAGAGAGGCTCTGCTGTGTCTGCTGCTGTGTGGAGCACT 2044
1997 CGCCACCGAGAAGCTGTGGGTGACCGTGTACTAGGCGGTGCCGTGTGGAAGGAGGCCAC 2056
2045 CTTCTGTTTCGCCAGCGCTAGCGAAACCCAGTCAACGGGGGAAGTGGCGGCCACACTGT 2104
2057 CACCACCTCTGCTGCGCCAGCGACCCAAAGGCTTACGACACCGAGGTGCACAACTGTG 2116
2105 GTCTGGAATTTGTAGCTCTCTGCAACAGCGGCGCAAGCAGAGCTCCAGTGTATCAACAC 2164
2117 GGCACCCACGCTGTGTCGCCACCGACCCCAACCCCGAGGAGTGGTCTGTGTAAGCT 2176
2165 CAAACGCGAGTGTGCACTCAATAGCACGCGCTGAACTGCAATGATGAGCTCAACACCGG 2224
2177 GACCGAGAATTCGACATGTGAAGAAACGACATGTTGGAGCAGATGCAACGAGGACATCAT 2236
2225 CTGGTTGGCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCTGAGAGGCT 2284
2237 CAGCCTGTGGGACACAGAGCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGAAGCT 2296
2285 AGCAGCTGCGACCCCTTACCGATTTTGACAGGGCTGGGGCCCTATCAGTTATGCCAA 2344
2297 GAAGTGACCGAGCTAGCACCAAGCTGCAACACCGCTGATCACCGAGCCTGCCCAA 2356
2345 CGGAAGCGCCCGGACAGCGGCCCTTACTGCTGGCACTACCCGCCAAAACCTTGGCGGTAT 2404
2357 GGTGAGCTTCGAGCCCATCCCAACACTACTGCGCCCGCGCGCTTCGCCATCATCTGAA 2416
2405 TGTGCCCGGAAGAGTGTGTGGTCCGGTATATGCTTCACTCCAGCGCCGTGTGTGT 2464
2417 GTGCAAGSACAAGAAGTTCAAGCGCAAGGGCCCTTCACCAACGTTGAGCACCGTGCAGTG 2476
2465 GGGAA----CGACCGAGCTGGGCGCGCCACCTACAGCTGGGGTGA---ATGA 2515
2477 CACCCAGCGATCCGCGCCGCTGTGAGCACCCAGCTGCTGCTGAAACGGTAGCTGCGCGA 2536
2516 TACGAGCTGTTGCTTTAAACAT--ACGAGGCCACCGCTGGGCAATGTTGGTGTGTA 2574
2537 GGAGGAGGTGGTGAATCCGAGCGCTAACTTCGCCGACAACGCCAAGGTGATCATCTGCA 2596

2575 CTTGGATGAACCTCACTGGATTCAACAAGTGTGGGAGCGCTCCTTGTGTCAATCGGAG 2634
2597 GCTGAACGAGAGCGTGGAGATCAACTGCAACCGCGCTAGCGCCCACTGCAACTGAGCCG 2656
2635 GGGCGGG-----CAACAACAACCTTGCACTGCCCACTGATTGCTTCCGCAA 2680
2657 CGCCAAGTGAAGACACACCTCTGAACAGATCGTGATCAAGCTGCGGAGCAGTTCCGCAA 2716
2681 GCATCCGAGCGCAACATACTCTCGGTGGGTCTCCGTCTCCGTGATC---ACACCCAGGTG 2737
2717 CAAGACCATCGTGTTCAGCACAGCAGCGCGCGGACCCCGAGATCGTGACCCACAGCTT 2776
2738 CTTGCTCGACTACCGGTATAGGCTTTGGCAATTATCCTTGTACCATCAACTACACCATATT 2797
2777 CAACTGCGCGCGAGTCTTCTTACTGCAACAGCGCTAGCTGCCCTGCGCATCAAGCA 2836
2798 TAAAAATCAGGATGTATACGTGGGAGGGTCAACACAGGCTGGAAGCTGCTGCAACTGGAC 2857
2837 GATCATCAACATGTGGCAGAAGTGGCAAGGCCATGTACGCCCCCCCATCAGCGGCCA 2896
2858 GCGGGCGGAACGTTGGATCTGGAA-----GATAGGAGCAGGTCCGA 2899
2897 GATCCGCTGCAGCAGCAACATCACCGGCTGCTGTGACCCCGCAGCGGCGCAACAGCAA 2956
2900 GATCGATATGGAGAAATCCTCAATATCCGAGAGTTCCTTAGGACCCCTGCTGTGTACAGCGGG 2959
2957 CAAAGAGCGAGATCTTCCGTCCGGCGCGCGGACATGCGGGAACAATGGGGCAGCA 3016
2960 GTTTTTCTTTGTGAACAAGAAATCCTCAATATCCGAGAGTCTAGACTCTGTGTGGTGAATTC 3019
3017 GCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGTGGCGTGGCCCCCACCAGGCCAA 3076
3020 TCTCAATTTTCTAG-----GGGATCTCCCGTGTGCTTTGGCCAAAATTCAGT 3069
3077 GCTTACCGTCCAGGCGCGCAGCTGCTGAGCGGCATCTGTGACGACGACAGCAACAACCTGCT 3136
3070 CCCAACCTCCAATCACTCACCAACCTCCTGCTCCCAATTTGCTGTGTTATCGCTGA 3129
3137 GCGCGCATGAGGCGCCAGCAGCACCTGCTGAGCTGACCGGTGGGGCATCAAGCAGCT 3196
3130 -----TGTGCTGCGCGCTTTTATCATATTCCTCTTTCATC----- 3164
3197 GCAGGCGCGCACCTTGGCGTGGAGCGCTACCTGAAGGACCAGCAGCAGCTCGAGCAGAT 3256
3165 CTGCTGCTATGCTCATCTTCTTATTTGTTCTTTGGAATATCAAGGTATGTTCCCGTT 3224
3257 TTGGAACCAACACCTTGGATGGAGTGGGACCGCGAGATCAACAACATACACAGCCTGAT 3316
3225 TGTCTCTTAATTTCCAG--GATCAACAACAACAGTACGGGACCATGCAAAAACCTGCACGA 3282
3317 CCACAGCTGATCGAGAGGCGCAGAACGACGAGAGAGAGAGGAGGAGGAGCTGTGGA 3376
3283 CTCCTGCTCAAG-----CAACTCTATGTTTCCCTCATGTTG 3319
3377 GCTGGACAGTGGCGCAGCTGTGGAACTGGTTCAACATCAACAACCTGCTGTGTGAGG 3436
3320 CTGTACAAAACCTACGATGGAAATTC---ACCTGTATTTCCATCCCATGCTCTCGGGC 3376
3437 ATCCAGATCTGCTGCTCTTCTAGTTGCGAGCATCTGTGTTGCGCCCTCCCGCTGCC 3496
3377 TTTTCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTTCT--CTTGGCTCAGTTTACTAG 3435
3497 TTCTTTGACCTGGAAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGGAATTCG 3556
3436 TGCCATTTGTTTCAGTGGTTCTGTAGGCTTTTCCCACTGCTTTTGGCTTTTCAGCTATATGA 3495
3557 ATCGCATTTCTCTGAGTAGGTGTCTTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAA 3616
3496 TGATGTGTTATTTGGGGCCAAAGTCTGTACAGCATCGTGTAGTCCCTTTATACCGCTTTAC 3555
3617 GGGGAGGATTTGGGAAGACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTATGGGTAC 3676
3556 CAATTTTCT-----TTTGTCTCTGGGTATACATTTAAGAAATTCAGACTC--- 3599

```
Db 3677 CCAGGTGCTGAAGAAATTGACCCGGTTCCTCTGGGCGAGAAAGACGAGCACATCCCT 3736
QY 3600 ---GAGCAAGTCTAGAAAGCGCGGCAAGATATCAAGGATCCACTAGCGGTAGAGCTCG 3656
Db 3737 TCTCTGTGACACACCTGTCTCAGCCCGCTGGTTCTTAGTTCAGAGCCCACTCATAGACA 3796
QY 3657 CTGATCAGCCTCGACTGTGCTTCTAGTTGCGAGCCATCTGTGTTGTCCTCCCGCT 3716
Db 3797 CTCATAGCTCAGGAGGGCTCCGCTTCAATCCACCGCTAAAGTACTTTGGAGCGGTCTC 3856
QY 3717 GCCTTCCTTGA-----CCCTGGAAGGTGCCATCTCCACCTGCTCTTCTTAATAAATGA 3770
Db 3857 TCCTCCCTCATCAGGCCCAACCAACCACTAGCCTCCAGAGTGGGAAGAAATTAAG 3916
QY 3771 GGAATTTGCATCGCATTTGCTGAGTAGTGTCTATTCTTCTGGGGGTGGGGTGGGCA 3830
Db 3917 CAAGATAGGCTATTAAGTCAGAGGAGGAGAAATGCTCCAACTGTGAGGAAGTAATG 3976
QY 3831 GGACAGCAAGGGGAGGATTTGGGAAGACAATAGCAGGCATGCTGG-GGAGCTCTCCGCT 3889
Db 3977 AGAGAAATCATAGAAATTTTAAGGCCATGATTTAAGGCCATCATGCGCTTAATCTTCGCT 4036
QY 3890 TCCTCGCTCACTGACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3949
Db 4037 TCCTCGCTCACTGACTCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4096
QY 3950 TCAAGCGCGTATACGGTTATCCAGAAATCAGGGGATAACCGAGGAAGACATGTGA 4009
Db 4097 TCAAGCGCGTATACGGTTATCCAGAAATCAGGGGATAACCGAGGAAGACATGTGA 4156
QY 4010 GCAAAAGGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGCGTGTGCTGGCGTTTCCAT 4069
Db 4157 GCAAAAGGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGCGTGTGCTGGCGTTTCCAT 4216
QY 4070 AGGCTCGGCGCGCTGACGAGCATCAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAAC 4129
Db 4217 AGGCTCGGCGCGCTGACGAGCATCAAAATTCAGCGCTCAAGTCAGAGGTGGCGAAAC 4276
QY 4130 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAGCTCCCTCGTGGCTCTCT 4189
Db 4277 CGACAGGACTATAAGATACAGGCGTTTCCCGCTGGAGCTCCCTCGTGGCTCTCT 4336
QY 4190 GTTCCGACCTGCGCTTACCGGATACCTGTCTCGCTTTCTCCCTTCGGGAAGCGTGGCG 4249
Db 4337 GTTCCGACCTGCGCTTACCGGATACCTGTCTCGCTTTCTCCCTTCGGGAAGCGTGGCG 4396
QY 4250 CTTTCTCAATGCTCAAGCTGATGATCTCAGTTTCGGGTGATGCTGCTTCCCTCAAGCTG 4309
Db 4397 CTTTCTCATAAGCTCAGCTGATGATCTCAGTTTCGGGTGATGCTGCTTCCCTCAAGCTG 4456
QY 4310 GGCTGTGTCAGCAACCCCGCTTACCGCGGCTGCGCTTATCCCGTAACTATCGT 4369
Db 4457 GGCTGTGTCAGCAACCCCGCTTACCGCGGCTGCGCTTATCCCGTAACTATCGT 4516
QY 4370 CTTGAGTCCAAACCGCTTAAGACACGACTTATCCGCACTGGCAGCAGCACTGTGTAACAG 4429
Db 4517 CTTGAGTCCAAACCGCTTAAGACACGACTTATCCGCACTGGCAGCAGCACTGTGTAACAG 4576
QY 4430 ATTAGCAGAGCGAGGTATGTAAGCGGTGTACAGATTCTTGAAGTGGTGGCTTAACCTAC 4489
Db 4577 ATTAGCAGAGCGAGGTATGTAAGCGGTGTACAGATTCTTGAAGTGGTGGCTTAACCTAC 4636
QY 4490 GGCTACACTAGAGGACAGTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4549
Db 4637 GGCTACACTAGAGGACAGTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4696
QY 4550 AAAAGAGTTGGTAGCTCTTGCATCCGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 4609
Db 4697 AAAAGAGTTGGTAGCTCTTGCATCCGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 4756
QY 4610 GTTTGCAAGCAGAGATTACGCGCAGAAAGAAAGGATCTCAAGAGATCTTTGATCTTT 4669
```

```
Db 4757 GTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTT 4816
QY 4670 TCTACGGGTCTGAGCGCTCAGTGGAAACGAAACTCAGCTTTAAGGGATTTTGGTCATGAGA 4729
Db 4817 TCTACGGGTCTGAGCGCTCAGTGGAAACGAAACTCAGCTTTAAGGGATTTTGGTCATGAGA 4876
QY 4730 TTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTAAATCAATC 4789
Db 4877 TTATCAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTAAATCAATC 4936
QY 4790 TAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGACCT 4849
Db 4937 TAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGACCT 4996
QY 4850 ATCTCAGCATCTGTCTATTTCCTTCATCCATGATTTGCCTGACTC 4894
Db 4997 ATCTCAGCATCTGTCTATTTCCTTCATCCATGATTTGCCTGACTC 5041

RESULT 13
US-10-359-120-51
; Sequence 51, Application US/10359120
; Publication No. US20040033487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKRABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; TITLE OF INVENTION: ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 6577
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
US-10-359-120-51

Query Match 38.8%; Score 2280.4; DB 17; Length 6577;
Best Local Similarity 68.8%; Pred. No. 0;
Matches 3627; Conservative 0; Mismatches 1151; Indels 496; Gaps 16;

QY 1 TCAGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAAGTCCCGGAGACGGTCA 60
Db 1 TCAGCGGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAAGTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGGTCAGCGGGTG 120
Db 61 CAGCTTGTCTGTAAAGCGGATGCGGAGCAGACAAAGCCGTCAGGCGGTCAGCGGGTG 120
QY 121 TTGGCGGGTGTCCGGGCTGGCTTAATCTATGCGGCATCAGACAGATTTGTACTCAGAGTGC 180
Db 121 TTGGCGGGTGTCCGGGCTGGCTTAATCTATGCGGCATCAGACAGATTTGTACTCAGAGTGC 180
QY 181 ACCATATGAAGCTTTTTCGAAAAGCCCTAGGCTCCAAAAAGCCCTCTCACTACTTCTGG 240
Db 181 ACCATATGCG-----TGTGAAATACCGCACA 207
QY 241 AATAGCTCAGAGCGCGGCGCTCGGCTCTGCTGATAAATAAAAAAATTAGTCAGCCA 300
Db 208 GATGCGTAAGGAGAAATATCCGCATCAG----- 235
```

QY	301	TGGGCGGAGNATGGCGGNACTGGGCGGAGGGAATATTGGCTATTGGCCATTGCAAT	360	1380	ATACTGTGCTTTCAGAGACTGACCGGACTGTGATATTTTACAGGATGGGT-CCATTT	1438
Db	236	-----ATTGGCTATTGGCCATTGCAAT	256	1277	ATACTGTGCTTTCAGAGACTGACCGGACTGTGATATTTTACAGGATGGGTGCCATTT	1336
QY	361	ACGTTGTAATCATATAATATGTAATTTATTTGGCTCATGTCCAATATGACCGCCA	420	1439	ATTATTTACAAATTCACATATACAAACGCGCTGCCCGTCCCGCAGTTTTTATATAA	1498
Db	257	ACGTTGTAATCATATAATATGTAATTTATTTGGCTCATGTCCAATATGACCGCCA	316	1337	ATTATTTACAAATTCACATATACAAACGCGCTGCCCGCAGTTTTTATATAA	1396
QY	421	TGTTGACATTGATTTAGACTAGTTAATAGTAATCAATTAACGGGGTCATTAGTTTCAT	480	1499	CATAGCGTGGGATCTCC---GACATCTCGGGTACGTTTCCGGACATGGGCTCTTCTCCG	1555
Db	317	TGTTGACATTGATTTAGACTAGTTAATAGTAATCAATTAACGGGGTCATTAGTTTCAT	376	1397	CATAGCGTGGGATCTCCACCGGAACTCCGGGTACGTTTCCGGACATGGGCTCTTCTCCG	1456
QY	481	AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCG	540	1556	GTAGCGGCGAGCTTCCACATCGAGCCTGTGTCATCGCTCCAGCGGCTCATGTGCG	1615
Db	377	AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCG	436	1457	GTAGCGGCGAGCTTCCACATCGAGCCTGTGTCATCGCTCCAGCGGCTCATGTGCG	1516
QY	541	CCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA	600	1616	TGGGAGCTCTCTTGTCTTAACAGTGGAGGCGCAGACTTAGGCACACACAATGCCACCA	1675
Db	437	CCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA	496	1517	TGGGAGCTCTCTTGTCTTAACAGTGGAGGCGCAGACTTAGGCACACACAATGCCACCA	1576
QY	601	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	660	1676	CCACAGTGTGCGCACAAGCGCTGGTAGGGTATGTCTGAAATGAGCTCGGAG	1735
Db	497	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA	556	1577	CCACAGTGTGCGCACAAGCGCTGGTAGGGTATGTCTGAAATGAGCTCGGAG	1636
QY	661	CATCAAGTATCATATGCAAGTCCGCCCTTATTGAGCTCAATGACCGTAAATGGCCC	720	1736	ATTGGGCTCGCACCTG-GACGCAGATGGAAGACTTAAGCAGCGGCAGAGAAGATGAC	1794
Db	557	CATCAAGTATCATATGCAAGTCCGCCCTTATTGAGCTCAATGACCGTAAATGGCCC	616	1637	ATTGGGCTCGCACCTG-GACGCAGATGGAAGACTTAAGCAGCGGCAGAGAAGATGAC	1696
QY	721	GCTGGCATATGCCAGTACATGACTTACGGGACTTTCTACTTTGGCAGTACATCTAC	780	1795	GCAGCTGAGTTGTTGATTTCTGATAAGAGTCAAGAGTAACTCCCGTTGCGGTGCTTTAA	1854
Db	617	GCTGGCATATGCCAGTACATGACTTATGGGACTTTCTACTTTGGCAGTACATCTAC	676	1697	GCAGCTGAGTTGTTGATTTCTGATAAGAGTCAAGAGTAACTCCCGTTGCGGTGCTTTAA	1756
QY	781	GTATTAGTCATCGCTATTACATGGTGTATGCGTTTGGCAGTACCAATGGCGGTGA	840	1855	CGGTGAGGCGAGTGTGATCTGAGCAGTACTCGTTGCTCGCGCGCGGCCACACAGACATA	1914
Db	677	GTATTAGTCATCGCTATTACATGGTGTATGCGTTTGGCAGTACCAATGGCGGTGA	736	1757	CGGTGAGGCGAGTGTGATCTGAGCAGTACTCGTTGCTCGCGCGCGGCCACACAGACATA	1816
QY	841	TAGCGTTTGGACTCAGCGGATTTCCAGTCTCCACCCCATGACGTCATGGGAGTTTG	900	1915	ATAGCTGACAGACTAACAGACTGTTCTTTCATGGGTCTTTTCTGACGTACCGTCTC	1974
Db	737	TAGCGTTTGGACTCAGCGGATTTCCAGTCTCCACCCCATGACGTCATGGGAGTTTG	796	1817	ATAGCTGACAGACTAACAGACTGTTCTTTCATGGGTCTTTTCTGACGTACCGTCTC	1876
QY	901	TTTTGGCACAAATCAACGGGACTTTCCAAATATGCTAAATAACCCCGCCGTTGACG	960	1975	GACGAATTCAGCA---ATCATGGATGCAATGAAGAGAGGCTCTGCTGTGCTGCTGC	2031
Db	797	TTTTGGCACAAATCAACGGGACTTTCCAAATATGCTAAATAACCCCGCCGTTGACG	856	1877	GACAGTGTGATCAGATATCGCGCGCTCTAGACACCATGCGGTGAAGGAGAGTACC	1936
QY	961	CAAAATGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTGTGTAAC	1020	2032	TGTGTGAGCAGACTCTTCTGTTTCGCGAGCGCTAGCGAAGAACCCACGTACCGGGGGAAGTG	2091
Db	857	CAAAATGGCGGTAGGCGTGTACGGTGGGAGTCTATATAAGCAGAGCTCGTTTGTGTAAC	916	1937	AGCACCTGTGGCGCTGGGCTGCGCTGGGGCACCATGTGCTGGGCATGCTGATATCT	1996
QY	1021	CGTCAGATCGCTGGAGAGCCCATCCACGTGTTTGTGACCTCCATAGAAGACACCGGAC	1080	2092	CGGGCCACACTGTGTGATTTGTAGCTCTCTCGCACCGCGCCAGCAGAACGTC	2151
Db	917	CGTCAGATCGCTGGAGAGCCCATCCACGTGTTTGTGACCTCCATAGAAGACACCGGAC	976	1997	GCAGCGCCACCGAGAGCTGTGGGTGACCGGTACTACGGGTGCGCGTGTGAGAGAGG	2056
QY	1081	CGATCCAGCTCCGCGCGGGAACGGTGCATTGGAAACGGGATCCCGTGCAGAGT	1140	2152	AGCTGATCAACACCAACCGCAGTTGGCACCCTCAATAGACACGG---CCCTGAACCTGCAATG	2208
Db	977	CGATCCAGCTCCGCGCGGGAACGGTGCATTGGAAACGGGATTCGCCGTGCAAGAGT	1036	2057	CCACCACCACTCTGCTCTCGCGCAGCGACGCAAGSCCTACGACACCGAGGTGCAACG	2116
QY	1141	GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1200	2209	ATAGCCTCAACACCGGCTGTGGCAGGCTTTTCTATCACCAAGTTCAACTTTCAG	2268
Db	1037	GACGTAAGTACCGCTTATAGACTCTATAGGCACACCCCTTTGGCTCTTATGCAATGCTATA	1096	2117	TGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCGCAGGAGGTGTGCTGTGA	2176
QY	1201	CTGTTTTTGGCTGGGCTCATACACCCCGG---TCCTTATGCTATAGGTGATGTTATGC	1259	2269	GTGTCTCTCAGAGGCTAGCCAGCTCGGACCCCTTACCGATTTTGTACAGGCTGGGGC	2328
Db	1097	CTGTTTTTGGCTGGGCTCATACACCCCGGTTCTTATGCTATAGGTGATGTTATGC	1156	2177	AGTGACCGAGAACTTTCGACATGTGGAAGAAAGACATGTGTGAGCAGATGACGAGACA	2236
QY	1260	TTAGCCTATAGGTGGGTATTGACATTAATGACACTCCCTTATTTGGTGACGATACT	1319	2329	CTATAGTTTATGCCAACGGAAGCGGCCCGACGAGCGCCCTACTGCTGGCACTACCCCC	2388
Db	1157	TTAGCCTATAGGTGGGTATTGACATTAATGACACTCCCTTATTTGGTGACGATACT	1216	2237	TCATCA----GCCTGTGGGACACAGGCTGAAGCCTCGGTGAAGCTGACCCCTGTG	2291
QY	1320	TTCCATTACTAATCCATAACATGGCTCTTTGGCACAACACTATCTATTTGGCTATATGCCA	1379	2389	CAAAACCTTGGCGTATTTGCGCGGGAAGAGTGTGTGGTCCGGTATATTGCTTCACTC	2448
Db	1217	TTCCATTACTAATCCATAACATGGCTCTTTGGCACAACACTATCTATTTGGCTATATGCCA	1276	2292	CGTGAGCTTGAAGTGACCGGACC-TGAAGAACGACACCAACCAACAGCAGCAGCGGCC	2350
				2449	CCAGCCCCGTGTGTGGTGGGAACGACCGACAGGTGCGG---CGCGCCCCACCTACAGCTGGG	2505

Db 2351 GCATGATCATGGAGAGGGCGAGATCAAGAACTGCGAGCTTCAACATCAGCACCGCATCC 2410
Qy 2506 GTGAAATGATAGCGAGCTCTCGTCTTAACTAACATACCGCCACCGCTGGGCAATTGGT 2565
Db 2411 GCGCAAGGTGCAAGAGGATGCGCTTCTTACAACTGGGACATCATCCCCATGACA 2470
Qy 2566 TCGGTTGTACCTGGATGAATCAACTGGAATTCACCAAGGTGTGGGAGCGCTCTCTGTG 2625
Db 2471 ACGACACCAAGCTACAGCTGACAGCTGCAACACCGCTAGCATCAATGCAACCGCC 2530
Qy 2626 TCATCGGAGGGGGGCAACAAACACCTGCACTGCGCCCACTGATGTGTTCCGCAAGCATC 2685
Db 2531 CCAACAAACACCGCAAGAGATCCACATCGCGCCCGCGCTTCTACACCAACCG 2590
Qy 2686 CGAGCGCCATATCTCTCGTGGGCTCCGCTCCCTGGATCACACCGAGTGCCTGGTGC 2745
Db 2591 GCGAGATCATCGGGACATCCGCGAGGCCCACTGCAACCTGAGCGCGCCCAAGTGAACG 2650
Qy 2746 ACTACCCGTATAGGCTTTGGCATTATCTTTGTATACCATCAACTACACCATATTTAAATCA 2805
Db 2651 ACACCTGAACAAGATCGTGATCAAGCTGCGCGAGAGTTCCGGCAACAAGACCATCGTGT 2710
Qy 2806 GGATGTAGTGGAGGGGTGAAACACAGGCTGGAAGCTGCTGCAACTGGAACGCGGGCG 2865
Db 2711 TCAAGCACAGACGCGGGCGGACCCCGAGATCGTGACCCACAGCTTCAACTGCGGGCGG 2770
Qy 2866 AACGTTGCGATCTGGAAGATAGGACAGGTCGAGATCGATATGGAGAACATCACATCAG 2925
Db 2771 AGTTCTTCTACTGCAACAGACCCAGCTGTTTCAACAGACCTGGTTCACAGACCTGGA 2830
Qy 2926 GATTCTTAGGACCCCTGCTGTTTACAGCGGGGTTTTCTTTGTGACAAGAAATCTCA 2985
Db 2831 GCACCGGGCAGCAACAAC-----ACGAGGGCAGCGACACCATCACCTGCGCTGCC 2884
Qy 2986 CAATACCGCAGATCTAGACTGTGGTGGATTCTCTCAATTTCTAGGGGATCTCCCG 3045
Db 2885 GCATCAAGCAGATCATCAACATGTGGCAGAAAGTGGGCAAGGCCATGTACGCCCCCCCA 2944
Qy 3046 TGTGTCTTGGCCAAATTCGAGTCCCACTCCCACTCCATCACTCAACCACTCTGTCTCTC 3105
Db 2945 TCAGCGGCCAGATCCCGTGCAGCAGCAACATCAACCGGCTGCTGAGCCGCGACGGCG 3004
Qy 3106 CAATTTGCTCTGTTATCGCTGATGTCTCGGGCGTTTTATCATATTTCTTCTATCC 3165
Db 3005 GCACAGCAACAGCAGAGCGATCTTCGTCGGGGCGGCGACATCGCGGCAACT 3064
Qy 3166 TGTGCTATGCTCATCTTCTTATGTTTCTTCTGATATC-----AAGGTATGTTGCC 3221
Db 3065 GCGCGAGCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTTGGGCGTGGCCCCCA 3124
Qy 3222 GTTTGCTCTAATTCAGGATCAACAAACAGTACGGGACCATGCAAAACCTGCAGC 3281
Db 3125 CCAAGGCCAAGCTTACCGTCCAGGCCGCGCAGCTGTGAGCGGCATCGTGACGACGAGA 3184
Qy 3282 ACTCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCTACGGATGA 3341
Db 3185 ACACCTGTGCGGCCATCGAGGCCGACGACGACCTGTCTGACGCTGACCGTGTGGGCA 3244
Qy 3342 AATTGACCTGTATTCCTATCCATCCATCGTCTCG-----GCTTTCGCAAAATACCTATGGAG 3398
Db 3245 TCAAGCAGCTGCAAGGCCGACCCCTGGCCGTGGAGCGCTACCTGAAGGACCGACGACG 3304
Qy 3399 TGGGCTTCAGTCCGTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAAGTGTTCGTA 3458
Db 3305 TCAGCAGATCTGGAACCAACACACCTGGATGAGTGGGACCGGAGATCAACAACATCA 3364
Qy 3459 GGGCTTTCCCCCACTGTTTGGCTTTCAGCTATATGATGATGTTGTTATGGGGGCCAAGT 3518
Db 3365 CAGCCTGATCCACAGCTGATCGAGGAGGACGACCAACGACGACGAGAACGACGAGG 3424
Qy 3519 CTGTACAGCATCGTGAAGTCCCTTTATACCGCTGTTTATACCAATTTCTTTTGTCTCTGGTA 3578

Db 3425 AGCTGCTGGAGCTGGCAAGTGGGCGAGCCCTGTGGAACCTGTTTCAACATCACCAACTGGC 3484
Qy 3579 TACATTTAAGAAATTCAGACTCGAGCAAGTCTAGAAAGGCGCGCAAGATATCAAGGATCC 3638
Db 3485 TGTGTGTACATCAAGCTGTTTCATCATGATCGTGGCGGCCCTGGTGGGCTCGGCATCGTGT 3544
Qy 3639 ACTACGCG--TTAGAGCTCGCTGATCAGCCCTCGACTGTGCTTCTTAGTTGCGAGCACTGTG 3697
Db 3545 TCGCCGTGTGAGCATCTGAGGATCCAGATCTGCTGTGCTCTTAGTTGCGAGCACTGTG 3604
Qy 3698 TTGTTTGGCCCTCCCGGCTGCTTCTTGAACCTCGAAGGTGCACTCCCACTGTCTCTTT 3757
Db 3605 TTGTTTGGCCCTCCCGGCTGCTTCTTGAACCTCGAAGGTGCACTCCCACTGTCTCTTT 3664
Qy 3758 CCTAATAAATGAGGAAATTCGATCGCATTTGCTGAGTGGTCTCATTTCTTCTGCGGG 3817
Db 3665 CCTAATAAATGAGGAAATTCGATCGCATTTGCTGAGTGGTCTCATTTCTTCTGCGGG 3724
Qy 3818 GTGGGTGGGCGAGCAGCAAGGGGAGGATTTGGGAAGACAAATAGCAGGCATGCTGGGG 3877
Db 3725 GTGGGTGGGCGAGCAGCAAGGGGAGGATTTGGGAAGACAAATAGCAGGCATGCTGGGG 3784
Qy 3878 AGC----- 3880
Db 3785 ATGCGGTGGGCTCTATGGGTACCCAGGTGCTGAAGAAATTGACCCGGTTCCTCTGGGCCA 3844
Qy 3881----- 3880
Db 3845 GAAAGAACGAGGCACATCCCTTCTCTGTGACACACCCTGTCCACGCCCCCTGGTCTTAG 3904
Qy 3881----- 3880
Db 3905 TTCCAGCCCCACTCATAGGACACTCATAGCTCAGGAGGCTCCGCTTCAATCCCACCCG 3964
Qy 3881----- 3880
Db 3965 CTAAAGTACTTGGAGCGGTCTCTCTCCTCATCAGCCCAACCAACCAACCTAGCCTC 4024
Qy 3881----- 3880
Db 4025 CAAGGTGGAAAGAAATTAAGCAAGATAGGCTATTAAAGTGCAGAGGAGAGAAATGCC 4084
Qy 3881----- 3880
Db 4085 TCCAACTGTAGGAGTAATGAGAGAAATCATAGAAATTTAAGGCCATCATGCGCTTAA 4144
Qy 3881 TCTTCGCTTCTCTGCTCATCTGACTGCTGCGCTCGGTGCTGCTGCGGCGAGCGGTA 3940
Db 4145 TCTTCGCTTCTCTGCTCATCTGACTGCTGCGCTCGGTGCTGCGGCGAGCGGTA 4204
Qy 3941 TCAGCTCACTCAAGGCGGTATACGGTTATCCACAGAAATCAGGGGATACGCGAGGAAG 4000
Db 4205 TCAGCTCACTCAAGGCGGTATACGGTTATCCACAGAAATCAGGGGATTAACGCGAGGAAG 4264
Qy 4001 AACATGTGACAAAAGGCCAGCAAAAGGCCAGAACCGTAAAGGCGCGCTTGTCTGGCG 4060
Db 4265 AACATGTGACAAAAGGCCAGCAAAAGGCCAGAACCGTAAAGGCGCGCTTGTCTGGCG 4324
Qy 4061 TTTTTCATAGGCTCGGCCCCCTGACGAGCATACAAAATTCGACGCTCAATGTCAGAG 4120
Db 4325 TTTTTCATAGGCTCGGCCCCCTGACGAGCATACAAAATTCGACGCTCAATGTCAGAG 4384
Qy 4121 TGGGNAACCCGACGAGCTATAAGATACAGGCGGTTTCCCTCGGAGCTCCCTCGTG 4180
Db 4385 TGGGNAACCCGACGAGCTATAAGATACAGGCGGTTTCCCTCGGAGCTCCCTCGTG 4444
Qy 4181 CGCTCTCTGTTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGA 4240
Db 4445 CGCTCTCTGTTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGA 4504
Qy 4241 AGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTGC 4300
Db 4505 AGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTGC 4564

QY 4301 TCCAAGCTGGCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCCTTATCCGGT 4360
Db |||||
QY 4361 AACTATCGTCTTGAGTCCACCCGGTAAGACACGACTTATCGCCACTGCGCAGCCACT 4420
Db |||||
QY 4421 GGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGG 4480
Db |||||
QY 4481 CCTAACTACGGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTGAGCCAGTT 4540
Db |||||
QY 4541 ACCTTCGGAAGAGTGTGTAGTCTCTGATCCGCCAACAACCAACCGCTGTAGCGGT 4600
Db |||||
QY 4601 GGTATCTTTTGTGCAAGCAGCAGATTACCGCGAGAAAAAGGATCTCAAGAAGATCCT 4660
Db |||||
QY 4661 TTGATCTTTTACGGGGTCTGACGCTCAGTGTGAACGAAAACTCAGGTTAAGGATTTTG 4720
Db |||||
QY 4721 GTCATGAGATTATCAAAAAGGATCTTCACTAGATCCTTTTAAATTAATAAATGAAGTTT 4780
Db |||||
QY 4781 AAATCAATCTAAAGTATATAGTAAACTTGTCTGACAGTTTACCAATGCTTAATCACT 4840
Db |||||
QY 4841 GAGGACCTATCTCAGCGATCTGTCTATTTGGTTTCAATAGTGTGCTGACTC 4894
Db |||||

RESULT 14
US-10-359-120-157
; Sequence 157, Application US/10359120
; Publication No. US2004003487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 6438
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
US-10-359-120-157

Query Match 38.7%; Score 2275.8; DB 17; Length 6438;
Best Local Similarity 69.5%; Pred. No. 0;
Matches 3558; Conservative 0; Mismatches 1232; Indels 333; Gaps 21;
QY 1 TCGCGGTTTCGGTGATGACGGTGAAAACTCTGCACATGACGCTCCCGGAGACGGTCA 60
Db |||||
QY 61 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAAGGCGGCTCAGCGGGTG 120
Db |||||
QY 121 TTGGCGGGTGTCCGGGCTGGCTTAATATGCGGCATCAGAGCAGATTGTACTGAGAGTGC 180
Db |||||
QY 181 ACCATATGAAGCTTTTTTGCAAAAGCTAGGCCTCCAAAAAGCCTCCTCACTACTTCTGG 240
Db |||||
QY 241 AATACTCAGAGCGCGGCGCCTCGGCTCTGCATTAATAAAAAATTAAGTCAGCCA 300
Db |||||
QY 301 TGGGGCGGAGATGGCGGAACTGGCGGGGAGGGAATTATTGGCTATTGGCCATTGCAT 360
Db |||||
QY 361 AGTTGTATCTATATCAATAATGTACATTTATATTGGCTCATGTCCAATATGACCGCCA 420
Db |||||
QY 421 TCTTGACATGATTAATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTCA 480
Db |||||
QY 481 AGCCCATATGCGAGTTCGCGGTACATAACTTACGTTAAATGGCCGCTGCTGACCG 540
Db |||||
QY 541 CCCAACGACCCCGCCCATTTGACGTCAATAATGAGCTATGTTCCCATAGTAACGCCAATA 600
Db |||||
QY 601 GGGACTTTCATGACGTCATGAGTGGGTGAGTATTTACGTTAACTGCGCACCTTGGCAGTA 660
Db |||||
QY 661 CATCAAGTGTATCATATGCCAGTCCGCCCTATTGAGTCAATGACGCTAAATGGCCC 720
Db |||||
QY 721 GCCTGGCATTTATGCCCAGTACATGACCTTACGGGACTTTTCTTACTTTGGCAGTACATCTAC 780
Db |||||
QY 781 GTATTAGTCATCGCTATTACATGTTGATGCGGTTTTTGGCAGTACACCAATGGGCGTGA 840
Db |||||
QY 841 TAGCGGTTTGAATCAGCGGATTTTCCAGTCTCCACCCCATTTGAGCTCAATGGGAGTTTG 900
Db |||||
QY 901 TTTTGGCACCAAAATCAACGGGACTTTTCCAAATGTGTAATAACCCGCCCTTGTGACG 960
Db |||||
QY 961 CAAATGGCGGTAGCGGTGTACGGTGGAGGTCTATATAGCAGAGCTGTTTAGTGAAC 1020
Db |||||
QY 1021 CGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTGTTTACCTCCATAGAAGACACCGGGAC 1080

QY 3137 GCGCGTCTTTATCATATTCCTCTTCAATCTCTGCTGCTATGCTCATCTTCTTATTTGGTTCT 3196
DB 3137 GGCACACAGATGAAGGACTGCACCGAGCGACAGGCTAAATTTTATGGGAAGATCTGGCT 3196
QY 3197 TCTGGATTATCAAGG-----TATGTTGCCCGTTTGTCTTAATTTCCAGGATCAACA 3248
DB 3197 TCCCACAAGGAAGGCCAGGGAATTTCTTTCAGAGCAGACGAGCAACAGCCCCACCA 3256
QY 3249 ACAACAGTACGGGACCATGCAAAACCTGCAAGACTCTCTGCTCA-----A 3293
DB 3257 GAAGAGACTTCAGGTTTGGGAAGAGACAACAACCTCCCTCTCAGAAAGCAGGAGCGGATA 3316
QY 3294 GGCAACTCTATGTTTCCCTCATGTTGCTGTGTAACAAAC----- 3330
DB 3317 GACAAGGAATCTGATCTTCTTAGCTTCCCTCAGATCACTCTTTGGCAGCAGCCCTCGTCA 3376
QY 3331 -----CTACGGATGGAATTCACCTGCTATTTCCATCCCATCGTCTGGGC 3376
DB 3377 CAATAAAGATAGGGGCCAGCTGAAGAGGCCCTTCTAGACAGGCCCTGGATCCAGATC 3436
QY 3377 TTTTCGCAAAATACCTATGGAGTGGGCCTCAGTCCGTTTCTCTTGCTCAGTT----- 3429
DB 3437 TGCCTGTCCTTCTAGTTGCCAGCATCTGTTGTTTGGCCCTCCCGTGCCTTCTCTTGAC 3496
QY 3430 -----TACTAGTGCCTATTTGTTTCAAGTGGTTCGTTAGGGCTTTTCCCCACTG 3474
DB 3497 CTTGGAAGGTGCCACTCCCACTGCTCTTCTTAATAAATGAGGAATTCATCGCATTTG 3556
QY 3475 TTTGGCTTTTCAGTATA-----TGCATGATGTTGTTTGGGGCCAAAGTCTGTACAGCA 3528
DB 3557 TCTGAGTAGGTGCATTTCTATTTCTGGGGGTGGGGTGGGGAGGACAGCAAGGGGAGGA 3616
QY 3529 TCGTAGTCCCTTTATACCGCTTTTACCAATTTTCTTTGCTCTCGGTATATATTTAAG 3588
DB 3617 TTGGGAAGACATAGCAGGCATGCTGGGNTGCGTGGGCTCTATGGTACCCAGTGTCT 3676
QY 3589 AA-----TTTCAGACTGAGCAAGTCTAGAAAGGCGCCCAAGATATCAAGGA 3635
DB 3677 GAAGAAATGACCCGGTTCTCTCGGCCAGAAAGAGCAGACATCCCTCTCTGTGA 3736
QY 3636 TCCACTACGGTTAGAGTCTGCTGATCAGCTGCTGCTCTTCTAGTTGCCAGCCATC 3695
DB 3737 CACACCTGTCCAGGCCCTGTTCTTAGTTTCAGCCCACTCATAGGACATCATAGCT 3796
QY 3696 TGTTGTTTGGCCCTCCCGCTGCTCTTTCAGCTTGAAGTGGAGTGCACCTCCCACTGCTCT 3755
DB 3797 CAGGAGGCTCCGCTTCAATCCACCCGCTAAAGTACTTTGGAGCGGTCTCTCCCTCCCT 3856
QY 3756 TTCTTAATAATGAGGAATTCGATCGCATTTGCTGAGTAGGTGTCATTTCTATTCGGG 3815
DB 3857 CATCAGCCCAACCAACCACTAGCTTCAAGAGTGGGAAGAAATTAAGCAAGATAGG 3916
QY 3816 GGGTGGGTGGGCAGACAGCA-----AGGGGAGGATTG 3851
DB 3917 CTATTAGTGCAGGGGAGAGAAATGCTTCCAAATGTGAGGAAGTAATGAGGAATC 3976
QY 3852 GGAAGACAATAGCAGGCATGCTGGGAGCTCTTCCGCTTCTCGCTCACTGACTCGCTGC 3911
DB 3977 ATAGAAATTTAAGGCCATCATGSCCTTAATCTTCCGCTTCTCGCTCACTGACTCGCTGC 4036
QY 3912 GCTCGTGTCTGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATACGTTAT 3971
DB 4037 GCTCGTGTCTGGCTCGGCGAGCGGTATCAGCTCACTCAAAAGGCGGTAAATACGTTAT 4096
QY 3972 CCACAGAACTCAGGGTAACGAGGAAGAAACATGTAGCAAAAGGCCAGCAAGGCCA 4031
DB 4097 CCACAGAACTCAGGGTAACGAGGAAGAAACATGTAGCAAAAGGCCAGCAAGGCCA 4156
QY 4032 GGAACCGTAAAGAGCGCGTCTGCTGGCGTTTTTCCATAGGCTTCCGCCCTCCCTGACGAGC 4091
DB 4157 GGAACCGTAAAGAGCGCGTCTGCTGGCGTTTTTCCATAGGCTTCCGCCCTCCCTGACGAGC 4216

QY 4092 ATCACAATAATCGACGCTCAAGTCAAGTCAAGTGGCGGAAACCCGACAGACTATATAAGATACC 4151
DB 4217 ATCACAATAATCGACGCTCAAGTCAAGTGGCGGAAACCCGACAGACTATATAAGATACC 4276
QY 4152 AGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTTCCGACCCCTGCCGCTTACCG 4211
DB 4277 AGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTTCCGACCCCTGCCGCTTACCG 4336
QY 4212 GATACCTGTCGCCCTTCTCCCTCGGGAAGCGTGGCGTCTTCTCAATGCTCAOCTGTGA 4271
DB 4337 GATACCTGTCGCCCTTCTCCCTCGGGAAGCGTGGCGTCTTCTCATAGCTCACGCTGTA 4396
QY 4272 GGTATCTCAGTTTCGGTGTAGTCTGCTCAAGCTGGCTGTGTGACAGAACCCCGCG 4331
DB 4397 GGTATCTCAGTTTCGGTGTAGTCTGCTCAAGCTGGCTGTGTGACAGAACCCCGCG 4456
QY 4332 TTCAGCCCCACCGCTGCGCTTATCCGGTAACTATCTGCTTTCAGTCCAAACCCGGTAAAGAC 4391
DB 4457 TTCAGCCCCACCGCTGCGCTTATCCGGTAACTATCTGCTTTCAGTCCAAACCCGGTAAAGAC 4516
QY 4392 AGCACTTATCGCACTGGCAGCAGCCACTGTGTAAACAGGATTAAGCAGAGCGGTATGTAG 4451
DB 4517 AGCACTTATCGCACTGGCAGCAGCCACTGTGTAAACAGGATTAAGCAGAGCGGTATGTAG 4576
QY 4452 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAAGACAGATAT 4511
DB 4577 GCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTTAACTAGAAAGACAGATAT 4636
QY 4512 TTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAAAGAGTTGCTAGTCTTGAT 4571
DB 4637 TTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAGAAAGAGTTGCTAGTCTTGAT 4696
QY 4572 CCGGCAAAACAAACCCGCTGCTAGCGGTGGTTTTTTTGTGTCAGAGCAGAGATTAACGC 4631
DB 4697 CCGGCAAAACAAACCCGCTGCTAGCGGTGGTTTTTTTGTGTCAGAGCAGAGATTAACGC 4756
QY 4632 GCAGAAAAAAGAGTCTCAAGAGATCTTTCATCTTTCTACGGGCTCTGACGCTCAGT 4691
DB 4757 GCAGAAAAAAGAGTCTCAAGAGATCTTTCATCTTTCTACGGGCTCTGACGCTCAGT 4816
QY 4692 GGAACGAAACTCAGCTTAAAGGATTTTGGTTCATGAGATTTCAAAAAGGATCTTCACTT 4751
DB 4817 GGAACGAAACTCAGCTTAAAGGATTTTGGTTCATGAGATTTCAAAAAGGATCTTCACTT 4876
QY 4752 AGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTT 4811
DB 4877 AGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTT 4936
QY 4812 GGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGACCTTATCTCAGCGATCTGTCTATTTC 4871
DB 4937 GGTCTGACAGTTTACCAATGCTTAAATCAGTGAGGACCTTATCTCAGCGATCTGTCTATTTC 4996
QY 4872 GTTTCATCCATAGTTGCTGACTC 4894
DB 4997 GTTTCATCCATAGTTGCTGACTC 5019

RESULT 15
US-10-359-120-32
; Sequence 32, Application US/10359120
; Publication No. US20040033487A1
; GENERAL INFORMATION:
; APPLICANT: NABLE, Gary J.
; APPLICANT: CHAKRABARTI, Bimal K.
; APPLICANT: HUANG, Yue
; TITLE OF INVENTION: MODIFICATIONS OF HIV Env, Gag, AND Pol
; TITLE OF INVENTION: ENHANCE IMMUNOGENICITY FOR GENETIC IMMUNIZATION
; FILE REFERENCE: NIH206.001C1
; CURRENT APPLICATION NUMBER: US/10/359,120
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: PCT/US01/25721
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/279,257

```
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/252,115
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/225,097
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 6577
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pVR1012x/s containing HIV genes
us-10-359-120-32

Query Match      38.7%; Score 2275.8; DB 17; Length 6577;
Best Local Similarity 69.2%; Pred. No. 0;
Matches 3664; Conservative 0; Mismatches 1097; Indels 530; Gaps 20;

QY 1 TC GCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAAGCTCCCGGAGACGGTCA 60
DB 1 TC GCGCGTTTCGGTGATGACGGTGAAACCTCTGACACATGCAAGCTCCCGGAGACGGTCA 60
QY 61 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAAGCCCGTCAAGCGGGTG 120
DB 61 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAAGCCCGTCAAGCGGGTG 120
QY 121 TTGGCGGGTGTCCGGGCTGGCTTAACCTATGCGGCATCAGACGAGATTGACTGAGAGTGC 180
DB 121 TTGGCGGGTGTCCGGGCTGGCTTAACCTATGCGGCATCAGACGAGATTGACTGAGAGTGC 180
QY 181 ACCATATGAAGCTTTTGGCAAAAGCCTTAGGCTCCCAAAAGCCTCCTCACTACTTCTGG 240
DB 181 ACCATATGCGG-----TGTAATAATCCGCACA 207
QY 241 AATAGCTCAGAGCGGAGGGGCTCGGCTCTGCAATAATAAAAAAATAGTCAGCCA 300
DB 208 GATCGCTAAGGAGAAAATACCGCATCAG-----235
QY 301 TGGGGCGGAGAAATGGCGGAACTGGGCGGAGGGAATTATTGGCTATTGGCCATTGGCAT 360
DB 236 -----ATTGGCTATTGGCCATTGGCAT 256
QY 361 ACCTTGTATCTATCATATAATATGATACATTTATATTGGCTCATGTCCCAATATGACCGCCA 420
DB 257 ACCTTGTATCTATCATATAATATGATACATTTATATTGGCTCATGTCCCAATATGACCGCCA 316
QY 421 TGTGACATTGATTTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTTCAT 480
DB 317 TGTGACATTGATTTATGACTAGTTATTAATAGTAATCAATTAACGGGGTCAATTAGTTTCAT 376
QY 481 AGCCCATATATGAGTTCCGCGTTACATACCTTACGTTAAATGGCCGCTGGCTGACCG 540
DB 377 AGCCCATATATGAGTTCCGCGTTACATACCTTACGTTAAATGGCCGCTGGCTGACCG 436
QY 541 CCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 600
DB 437 CCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAAGCCCAATA 496
QY 601 GGGACCTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCGCCAATTGGGAGTA 660
DB 497 GGGACCTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCGCCAATTGGGAGTA 556
QY 661 CATCAAGTGTATCATATGCAAGTCCGCCCTTATTGAGTCAATGACGGTAAATGGCCC 720
DB 557 CATCAAGTGTATCATATGCAAGTACGCCCTTATTGAGTCAATGACGGTAAATGGCCC 616
QY 721 GCCTGGCATTTATGCCAGTACATGACCTTACGGGACTTTCTACTTTGGGAGTACATCTAC 780
DB 617 GCCTGGCATTTATGCCAGTACATGACCTTATGGGACTTTCTACTTTGGGAGTACATCTAC 676
QY 781 GTATTAGTCATCGCTATTACATGGTGTGCGGTTTGGCAGTACCAATGGCGGTGGA 840
DB |||||
```

1915 A TAGCTGACAGACTAACAGACTGTTCCCTTTCATGGGTCTTTTCTGCACTACCGTCGTC 1974
Db |||||
1817 ATAGCTGACAGACTAACAGACTGTTCCCTTTCATGGGTCTTTTCTGCACTACCGTCGTC 1876
QY |||||
1975 GACGAATTCAAGCAATCATGGATG-----CAATGAAGAGAGGCTCTGCTGTGCTGT 2029
Db |||||
1877 GACACGTGTGATCAGATATCGCGGCGCTCTAGAAATGCGCGTGAAGAGAGATACGACA 1936
QY |||||
2030 GCTGTGTGAGCAGTCTTCTGTTTCGCCAGCGCTAGCGAAACCCACGTCACCGGGGAAG 2089
Db |||||
1937 CCTGTGCGCTGGGCTGGCGCTGGGCGACCACTGCTGCTGGGCATCTGATGATCTGCAG 1996
QY |||||
2090 TGGCGGCGACACTGTGTCTGGAATTTGTAGCTCTCTCGACACAGGCGCCAGACGAAGCT 2149
Db |||||
1997 CGCCACCGAAGCTGTGGGTGACCGCTGACTACGGCGTGCCCGTGTGGAAGAGGCCAC 2056
QY |||||
2150 CCAGCTGATCAACACCAAGCGAGTTGGGACCTCAATAGCAGCGCCCTGAACCTGCAATGA 2209
Db |||||
2057 CACCACTTGCTCTGCGCCAGCGACGCCAAGGCTTAGCACCGAGGTGCA---CAACGT 2113
QY |||||
2210 TAGCTTCAACACCGGCTGGTGGCAGGCTTTTCTATCACCAAGTTTCAAATCTTTTCAG 2269
Db |||||
2114 GTGGCCACCCAGCGCTGCGTCCACCGACCCCAACCCCGAGGAGTGTGCTGTGAA 2173
QY |||||
2270 CTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGAATTTTGAACAGGCTGGGGCCC 2329
Db |||||
2174 CGTGACCGAAGCTTCGCATGTGTGAAGAAACGACATGCTGTGGAGCAGATGACACGAGCAT 2233
QY |||||
2330 TATCAGTTATGCCAAGCGAGCGGCCCGACGAGCGCCCTACTGCTGGCAGCTACCCCCC 2389
Db |||||
2234 CATCA-----GCTGTGGACACAGAGCTGAAGCCCTGCGTAGCTGACCCCTGCTGTC 2288
QY |||||
2390 AAAACCTTGGGTATTTGTCGCCGGAAGAGTGTGTGTGTCGGGTATATTTGCTTCACTCC 2449
Db |||||
2289 GTGAGCCTGAAGTGCAACG-ACGCTAGCAAGAACTGCAGCTTCAACATCAGCACCAAGCAT 2347
QY |||||
2450 CAGCCCGTGTGTGGGAAGAACGACGACAGTGTGGCGCGGCCCACTACAGCTGGGGTGA 2509
Db |||||
2348 CCGCGGCAAGGTGCAGAAGAGTACGCCCTTCTTACAAGCTGGACATCATCCCCATCGA 2407
QY |||||
2510 AAATGATACGGAGCTCTGCTCCTTAAACAATACGAGCCACGCTGGGCAATTTGGTTGG 2569
Db |||||
2408 CAAAGACACACAGCTACAGCTGACAGCTGACAGTGCACACACAGGCTGATCACCGAG--GC 2464
QY |||||
2570 TTGTACTTGGATGAATCAACTGGAATTTACCAAGTGTGGGAGCGCTCTTTGTGTGAT 2629
Db |||||
2465 CTGCCCAAGGTGAGCTTCGAGGCCATCCCAACCACTACTGCGCCCCCGCGGCTTGC 2524
QY |||||
2630 CGGAGGGCGGGCAACAAACCTGCACTGCCCCCACTGATTTGTTCCGCAAGCATCCGGA 2689
Db |||||
2525 CATCTGAAAGTGCAGGACAAAGATTCAACGGCAAGGGCCCCCTGCACCAACGTGAGCAC 2584
QY |||||
2690 CGCCACATCTCTCGGTGGGCTCCGGTCCCTGGATCACAC-----CCAGGTG 2737
Db |||||
2585 CGTGCAGTGACCCACGGCATCTCGCCCCCGTGTGTGAGCACCCAGCTGCTGCTGAACGGTAG 2644
QY |||||
2738 CTTGTGTGACTACCCCTATAGGCTTTGGCAATTATCTTGTGTACCATCAACTACACCAATT 2797
Db |||||
2645 CTTGGCCGAGGAGGAGTGTGTATCCGAGCGCTAACTTTCGCCGCAACCGCAAGGTGAT 2704
QY |||||
2798 TAAATCAGATGTAGCTGGAGGGGTGAAACACAGCTGTGAAGCTGCTGCAACTGGAC 2857
Db |||||
2705 CATCTGTGACGTGAACGAGAGCGTGGAGATCAACTGCACCGCGCTAGCGGCCCACTGCAA 2764
QY |||||
2858 GCGGGCGGAAGCTTGGCATCTGAAGATAGGACAGGTCGAGATCGATATG-GAGAAACA 2916
Db |||||
2765 CTTGAGCCGGCCAACTGGAAACGACACCTGTGAACAGATCTGTGATCAAGCTGGCGAGCA 2824
QY |||||
2917 TCACATCAGATTTCTAGGACCCCTCTGTGTGTACAGCGGGGTTTTTTCTTTGTGACAA 2976
Db |||||
2825 GTTCGGCAACAAGACCATCGTGTGTCAAGCACAGCAGCGCGCGGCGACCCCGAGATCTGTAC 2884

2977 GAATCTCACAATACGCAGAGCTTAGACTCGTGGTGAGTCTCTCAATTTCTTAGGGG 3036
Db |||||
2885 CCACAGCTTCAACTGCGGCGGAGTTCTTCTACTGCAACAGCGCTAGCTGCGCTGCGG 2944
QY |||||
3037 GATCTCCCGTGTCTTGGCCAAAAATTCGCACTTCCCAACCTCCCAATCACTCACCACCT 3096
Db |||||
2945 CATCAAGCAGATCA-TCAACATGTGGCAAGAGTGGCAAGCCATGTAGCCCCCCCCCA 3003
QY |||||
3097 CTTGCTCTCCAAATTTGCTGCTGCTGATTCGCTGATGTCTGCGGCGTTTATCATATTC 3156
Db |||||
3004 TCAGCGGCAGATCGCTGACAGCAACATCACCGGCTGCTGTGACCCGCGAGCGG 3063
QY |||||
3157 TCTTCATCTGCTGCTATGCTCTCATCTTCTTATTTGTTCTTCTGATTTATCAAGGTATGT 3216
Db |||||
3064 GCAACAGCAACAAACGAGAGCGAGATCTCCGTCGCGGCGCGCGACATGCGGCACACT 3123
QY |||||
3217 TGC-CCGTTTGTCTCTTAATTCAGGATCAACAAACAGTACGGGACCATGCAAAACC 3275
Db |||||
3124 GCGCAGCGAGCTGTACAGTACAGGTGGTGAAGATCGAGCCC-----CTGGGCG 3174
QY |||||
3276 TGCACGACTCTGCTCAAGGCAACTATATGTTTCCCTCATGTTGCTGTATCAAAACCTACG 3335
Db |||||
3175 TGGCCCCCACCACCAAGGCCAAGCTTACCGTCCAGGCGCGCAGCTGCTGAGCGGCATCGTC 3234
QY |||||
3336 GATGGAATTTGACCTGTATTTCCCATCCATCGTCTGCGCTTTCGCAAAATACCTATGG 3395
Db |||||
3235 AGCAGCAGAAACCTGCTGCGCGCATCGAGGCCCAGCAGCACTGCTGCACTGACCG 3294
QY |||||
3396 GAGTGGGCTCAGTCCGTTTCTTTGGCTCAGTTACTAGTGCCTTTTCTCAGTGGTTTC 3455
Db |||||
3295 TGTGGGCGCATCAAGCAGCTGCAGG-CCCGCACCTTGGCGCGCTGAGCGCTACCTGAAGGAC 3353
QY |||||
3456 GTAGGGCTTTCCCCACTGTTTGGCTTTTCAGTATATGATGATGATGTTGTTGGGGGCA 3515
Db |||||
3354 CAGCAGCAGCTCGAGCAGATCTGGAACACACACCACTGATGAGTGGGACCGCAGATC 3413
QY |||||
3516 AGTCTGTACAGCATCGTGTGCTCCCTTATACCGCTGTTACCAATTTCTTTTGTCTCTGG 3575
Db |||||
3414 AACAATCACAGCCTGATCCAC-----ACGCTGATCGAGGAGCGCAGAACCGAG 3467
QY |||||
3576 GTATACATTTAAGAAATTCAGACTCGAGCAAGTCTAGAAAGCGCGCAAGATATCAAGGA 3635
Db |||||
3468 GAGAAGACGACGAGGAGCTGCTGGAGCTGGACAGTGGGCCAGCTGTGGAATCGTTTC 3527
QY |||||
3636 TCCACTACCGTTGAGCT--CGCTGATCAGCGCTGCACTGTGCTTCTTAGTTGCGCAGCA 3693
Db |||||
3528 AACATCACCAACTGGCTGTGTGAGGATCCAGATCTGCTGTGCTTCTTAGTTGCCAGCA 3587
QY |||||
3694 TCTGTGTTTGGCCCTCCCGTGTCTTGTAGCCCTGGAAGGTGCCACTCCCACTGTC 3753
Db |||||
3588 TCTGTGTTTGGCCCTCCCGTGTCTTGTAGCCCTGGAAGGTGCCACTCCCACTGTC 3647
QY |||||
3754 CTTTCTTAATAAATGAGGAAATTCATCGCATTTGCTGAGTAGTGTCTATTTCTTCTG 3813
Db |||||
3648 CTTTCTTAATAAATGAGGAAATTCATCGCATTTGCTGAGTAGTGTCTATTTCTTCTG 3707
QY |||||
3814 GGGGTGTGGGTGGGCGAGACAGCAAGGGGAGGATTTGGGAAGAACAATAGCAGGCAATGCT 3873
Db |||||
3708 GGGGTGTGGGTGGGCGAGCAGCAAGGGGAGGATTTGGGAAGACAAATAGCAGGCAATGCT 3767
QY |||||
3874 GGGGAGC----- 3880
Db |||||
3768 GGGGTGTGGGTGGGCTCTATGGGTACCCAGGTGCTGAAGAAATTGACCCCGTTCTCTCTGG 3827
QY |||||
3881 ----- 3880
Db |||||
3828 GCCAGAAAGAACGAGCAGCAGCATCCCTTCTCTCTGTGACACACCTGTGCCAGCCCTGTTTC 3887
QY |||||
3881 ----- 3880
Db |||||
3888 TTAGTTCCAGCCCCACTCATAGGACACTCATAGCTCAGGAGGCTCCGCTTCAATCCCA 3947
QY |||||
3881 ----- 3880

```
Db 3948 CCGCTAAAGTACTTGGAGCGGTCTCTCCCTCCCTCATCAGCCCAACCAACCTAG 4007
Qy 3881 ----- 3880
Db 4008 CTTCCAAAGTGGGAAGAAATTAAGCAAGATAGGCTATTAAAGTGCAGAGGGAGAGAAA 4067
Qy 3881 ----- 3880
Db 4068 TGCCTCCAACATGTGAGGAAGTAATGAGAGAAATCATAGAAATTTAAAGGCCATGATTTAA 4127
Qy 3881 ----- 3880
Db 4128 GGCATCATGCGCTTAAATCTTCGGCTTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTTC 4187
Qy 3924 GGCCTGGCGGAGCGGTATCAGCTCACTCAAGCGGTAAACCGTTATCCACAGAAATCAG 3983
Db 4188 GGCCTGGCGGAGCGGTATCAGCTCACTCAAGCGGTAAACCGTTATCCACAGAAATCAG 4247
Qy 3984 GGGATAACGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAA 4043
Db 4248 GGGATAACGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAA 4307
Qy 4044 AGGCCGCGTTCGTGGCGTTTTTCCATAGGCTCGCGCCCCCTGAGAGCATCAAAAAATC 4103
Db 4308 AGGCCGCGTTCGTGGCGTTTTTCCATAGGCTCGCGCCCCCTGAGAGCATCAAAAAATC 4367
Qy 4104 GAGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATACCAAGCGGTTTCCCC 4163
Db 4368 GAGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATACCAAGCGGTTTCCCC 4427
Qy 4164 CTGGAAGCTCCCTCGTGGCGTCTCTCTGTTCCGACCTTACCGGATACCTGTCCG 4223
Db 4428 CTGGAAGCTCCCTCGTGGCGTCTCTCTGTTCCGACCTTACCGGATACCTGTCCG 4487
Qy 4224 CTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCAATGCTCACGCTGAGGTATCTCAGTT 4283
Db 4488 CTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCAATGCTCACGCTGAGGTATCTCAGTT 4547
Qy 4284 CGGTGTAGTGTCTCGCTCAAGCTGGGCTGTGCAAGAACCCCGTTTCAGCCGACC 4343
Db 4548 CGGTGTAGTGTCTCGCTCAAGCTGGGCTGTGCAAGAACCCCGTTTCAGCCGACC 4607
Qy 4344 GCTGCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCCGTTAAGACAGACTTATCGC 4403
Db 4608 GCTGCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCCGTTAAGACAGACTTATCGC 4667
Qy 4404 CACTGCAGCAGCACTGGTAAAGGATAGCAGAGCGAGGTATGAGCGGTGCTACAG 4463
Db 4668 CACTGCAGCAGCACTGGTAAAGGATAGCAGAGCGAGGTATGAGCGGTGCTACAG 4727
Qy 4464 AGTTCTTGAAGTGTGGCTTAAGTACGGCTACACTAGAGGACAGTATTGGTATCTCGG 4523
Db 4728 AGTTCTTGAAGTGTGGCTTAAGTACGGCTACACTAGAGGACAGTATTGGTATCTCGG 4787
Qy 4524 CTCTGCTGAAGCCAGTTACCTTCGGAAGAGGTGGTAGTCTCTGATCCGCAACAAA 4583
Db 4788 CTCTGCTGAAGCCAGTTACCTTCGGAAGAGGTGGTAGTCTCTGATCCGCAACAAA 4847
Qy 4584 CCACCGCTGGTAGCGGTGGTTTTTGTGTCAGGAGCAGAGATTACCGCAGAAAAAAG 4643
Db 4848 CCACCGCTGGTAGCGGTGGTTTTTGTGTCAGGAGCAGAGATTACCGCAGAAAAAAG 4907
Qy 4644 GATCTCAAGAGATCTTTGATCTTTTACGGGCTGACGCTCAGTGGAGCAAAACT 4703
Db 4908 GATCTCAAGAGATCTTTGATCTTTTACGGGCTGACGCTCAGTGGAGCAAAACT 4967
Qy 4704 CACGTTAAGGGATTTTGGTCAAGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAA 4763
Db 4968 CACGTTAAGGGATTTTGGTCAAGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAA 5027
Qy 4764 ATTAAAAATGAAGTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT 4823
```

```
Db 5028 ATTAAAAATGAAGTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT 5087
Qy 4824 ACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGCTATTCTTTCGTTTCATCCATAG 4883
Db 5088 ACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGCTATTCTTTCGTTTCATCCATAG 5147
Qy 4884 TTGCCTGACTC 4894
Db 5148 TTGCCTGACTC 5158
```

Search completed: April 30, 2005, 17:03:54
Job time : 3237 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2005, 21:50:43 ; Search time 16330 Seconds
(without alignments)

13710.596 Million cell updates/sec

Title: US-10-715-665-6

Perfect score: 5882

Sequence: 1 tcgcgcgttcggtgatgac.....tatcacgagcccttcgcg 5882

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1052.6	17.9	1070	1	AJ281552
2	1015	17.3	1048	7	CO552396
3	986.4	16.8	1013	4	BM438846
4	920.8	15.7	1004	1	AJ281480
5	886.6	15.1	928	7	CO487414
6	876.4	14.9	902	7	CR753463
7	875.8	14.9	935	4	BG838279
8	841	14.3	841	1	AL042026
9	824.6	14.0	854	4	BM438950
10	817	13.9	865	7	CK125894
11	814.4	13.8	1126	8	BZ577702
12	809.4	13.8	856	7	CN823189
13	808	13.7	819	6	CD649375
14	797.6	13.6	1011	8	BZ576726
15	795.6	13.5	820	7	CR753457
16	793	13.5	827	7	CN823902
17	791.8	13.5	806	4	BJ684280
18	782.2	13.3	1574	8	BZ572566
19	774.8	13.2	1336	8	BZ575810
20	774.6	13.2	800	1	AJ281449
21	774.4	13.2	789	6	CD280920
22	773.2	13.1	1370	8	BZ571721
23	772.6	13.1	1073	7	CF269652
24	771.4	13.1	832	7	CN822433

25	766.4	13.0	846	7	CV468077
C 26	765.8	13.0	769	7	CV224987
C 27	763.2	13.0	780	5	BQ825693
C 28	763.2	13.0	1249	8	BZ572284
C 29	758.2	12.9	966	8	BZ570738
C 30	757	12.9	759	6	CD279661
C 31	755.4	12.8	1463	8	BZ571475
C 32	751.4	12.8	786	7	CN823164
C 33	749	12.7	752	7	CR766850
C 34	743.8	12.6	966	8	BZ575002
C 35	743.6	12.6	832	4	BG923768
C 36	743.4	12.6	1003	8	BZ576686
C 37	742.2	12.6	844	8	BZ574513
C 38	741.4	12.6	758	9	CL422788
C 39	741	12.6	741	6	CD279174
C 40	738	12.5	914	8	BZ569398
C 41	734.4	12.5	802	7	CN823157
C 42	733.8	12.5	950	8	BZ571129
C 43	732.4	12.5	998	8	BZ576702
C 44	730	12.4	730	6	CD281097
C 45	730	12.4	811	9	ATH517156

ALIGNMENTS

RESULT 1
AJ281552 1070 bp mRNA linear EST 30-JUN-2000
LOCUS 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION AJ281552
VERSION AJ281552.1 GI:6929432
KEYWORDS EST
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Borik,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
PUBMED 10841561
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/z"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match		17.9%;	Score 1052.6;	DB 1;	Length 1070;
Best Local Similarity		99.5%;	Pred. No. 5.5e-309;		
Matches 1066;		Conservative	0;	Mismatches	4; Indels 1; Gaps 1;
QY	4052	TTGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCA	4111		
Db	1	TTGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGACGCTCA	60		
QY	4112	AGTCAGAGGTGGGAAACCCGACAGACTATAAAGATACACAGCGTTTCCCGCTGGAAGC	4171		
Db	61	AGTCAGAGGTGGGAAACCCGACAGACTATAAAGATACACAGCGTTTCCCGCTGGAAGC	120		
QY	4172	TCCCTCGTGGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTCTC	4231		
Db	121	TCCCTCGTGGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCGCTTCTC	180		
QY	4232	CTTTCGGGAAGCGTGGCGTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTAG	4291		
Db	181	CTTTCGGGAAGCGTGGCGTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGGTAG	240		
QY	4292	GTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGCTGCGCC	4351		
Db	241	GTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGCTGCGCC	300		
QY	4352	TTATCCGGTAACTATGCTTGTAGTCCAAACCGGTAAAGACAGCACTTATCGCCACTGGCA	4411		
Db	301	TTATCCGGTAACTATGCTTGTAGTCCAAACCGGTAAAGACAGCACTTATCGCCACTGGCA	360		
QY	4412	GCAGCCACTGGTAAACAGGATTAAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTG	4471		
Db	361	GCAGCCACTGGTAAACAGGATTAAGCAGCGAGGTATGTAGGCGGTGCTACAGATTCTTG	420		
QY	4472	AAGTGTGGCTTAACCTACCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG	4531		
Db	421	AAGTGTGGCTTAACCTACCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG	480		
QY	4532	AAGCCAGTTACCTTCGGAAGAGTTGTAGTCTTGTATCGGCAACAAACACCCGCT	4591		
Db	481	AAGCCAGTTACCTTCGGAAGAGTTGTAGTCTTGTATCGGCAACAAACACCCGCT	540		
QY	4592	GGTAGCGGTGGTTTTTGTGTTGCAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAA	4651		
Db	541	GGTAGCGGTGGTTTTTGTGTTGCAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAA	600		
QY	4652	GAAGATCTTTTATCTTTCTACGGGGTCTGACGCTCAGTGGAAAGAAATCAAGTTAA	4711		
Db	601	GAAGATCTTTTATCTTTCTACGGGGTCTGACGCTCAGTGGAAAGAAATCAAGTTAA	660		
QY	4712	GGGATTTTGGTCATGAGATTATCAAAAGAGATCTTACCTAGATCTCTTTTAAATTA	4771		
Db	661	GGGATTTTGGTCATGAGATTATCAAAAGAGATCTTACCTAGATCTCTTTTAAATTA	720		
QY	4772	TGAAGTTTAAATCAATCTAAAGATATATAGTAAACTTGGTCTGACAGTTTACCAATGC	4831		
Db	721	TGAAGTTTAAATCAATCTAAAGATATATAGTAAACTTGGTCTGACAGTTTACCAATGC	780		
QY	4832	TTAATCAGTGAAGCACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTGCCTGA	4891		
Db	781	TTAATCAGTGAAGCACTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTGCCTGA	840		
QY	4892	CTCCCGCTCGTGTAGTAATACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCA	4951		
Db	841	CTCCCGCTCGTGTAGTAATACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCA	900		
QY	4952	ATGATACCGGAGACCCAGCTCACCGGCTCCAGATTATACAGCAATAAACAGCCAGCC	5011		
Db	901	ATGATACCGGAGACCCAGCTCACCGGCTCCAGATTATACAGCAATAAACAGCCAGCC	960		
QY	5012	GGAAGGGCCGAGCAGAAAGTGGTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAT	5071		
Db	961	GGAAGGGCCGAGCAGAAAGTGGTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAT	1019		
QY	5072	TGTTTCGGGGAAGCTAGATAGTAGTTCGCCAGTTAATAGTTTGGCGCAAC	5122		

Db	1020	TGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGCAAC	1070		
----	------	--	------	--	--

RESULT 2

LOCUS

CO552396

DEFINITION

ACly4_50 Sea lamprey AcLy Petromyzon marinus cDNA, mRNA sequence.

ACCESSION

CO552396

VERSION

CO552396.1

GI:51800732

KEYWORDS

EST.

SOURCE

Petromyzon marinus (sea lamprey)

ORGANISM

Petromyzon marinus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; Petromyzontiformes; Petromyzontidae; Petromyzon.

REFERENCE

1 (bases 1 to 1048)

AUTHORS

Pancer, Z., Mayer, W.E., Klein, J. and Cooper, M.D.

TITLE

Prototypic T-cell receptor and CD4-like coreceptor expressed in lymphocytes of the agnathan sea lamprey

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 101, 13273-13278 (2004)

COMMENT

Contact: Pancer, Zeev
Division of Developmental and Clinical Immunology
The University of Alabama at Birmingham
378 Wallace Tumor Institute, 1530 Third Avenue, South, Birmingham, AL 35294-3300
Tel: 205-975-5812
Fax: 205-975-7218
Email: zpancer@uab.edu.

FEATURES

source

1..1048

Location/Qualifiers

/organism="Petromyzon marinus"

/mol_type="mRNA"

/db_xref="taxon:7757"

/cell_type="lymphocyte"

/dev_stage="immune stimulated larvae"

/clone_lib="Sea lamprey Acly"

/note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR subtracted cDNA libraries of immune stimulated larvae. All are single pass 5' or 3' sequences randomly cloned in pGEM-T Easy (Promega)."

ORIGIN

Query Match

17.3%;

Score 1015;

DB 7;

Length 1048;

Best Local Similarity

99.0%;

Pred. No. 1.7e-297;

Matches 1041;

Conservative

1;

Mismatches

6;

Indels

3;

Gaps

2;

QY

3988

TAAACAGGAAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGCAACCGTAAAAAGGC

4047

Db

1

TAAACAGGAAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGCAACCGTAAAAAGGC

60

QY

4048

CGCGTTGCTGGCGTTTTCATAGGCTCGCGCCCTGACGAGCATCAAAAATCGACG

4107

Db

61

CGCGTTGCTGGCGTTTTCATAGGCTCGG -CCCGCTGACGAGCATCAAAAATCGACG

119

QY

4108

CTCAAGTCAGAGTCGGCAAAACCCGACAGGACTATAAAGATACCGAGCGTTTCCCGCTGG

4167

Db

120

CTCAAGTCAGAGTCGGCAAAACCCGACAGGACTATAAAGATACCGAGCGTTTCCCGCTGG

179

QY

4168

AAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCCGCCCTT

4227

Db

180

AAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCCGCCCTT

239

QY

4228

TCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAAGCTAGGTATCTCAGTTTCGGT

4287

Db

240

TCTCCCTTCGGGAAGSGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTTCGGT

299

QY

4288

GTAGTTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTG

4347

Db

300

GTAGTTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCGACCGCTG

359

QY

4348

CGCCTTATCCGGTAACTATGCTCTTGTAGTCCAAACCCCGGTAAAGACAGCACTTATCGCACT

4407

Db

360

CGCCTTATCCGGTAACTATGCTCTTGTAGTCCAAACCCCGGTAAAGACAGCACTTATCGCACT

419


```
QY 4408 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATCTAGCGGTGCTACAGAGTT 4467
Db 420 GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATCTAGCGGTGCTACAGAGTT 479
QY 4468 CTTGAAGTGTGGCTTAACCTACCGCTACCTAGAGGACAGTATTTGGTATCTCGGCTCT 4527
Db 480 CTTGAAGTGTGGCTTAACCTACCGCTACCTAGAGGACAGTATTTGGTATCTCGGCTCT 539
QY 4528 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGGATCCGGCAAAACCAACCAC 4587
Db 540 GCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGGATCCGGCAAAACCAACCAC 599
QY 4588 CGCTGGTAGCGGTGTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATC 4647
Db 600 CGCTGGTAGCG--TGTTTTTGGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATC 657
QY 4648 TCAAGAAGATCCTTTGATCTTTTCTAGCGGGTCTGACGCTCAGTGGAAACGAAACTCACG 4707
Db 658 TCAAGAAGATCCTTTGATCTTTTCTAGCGGGTCTGACGCTCAGTGGAAACGAAACTCACG 717
QY 4708 TTAAGGATTTTGGTCAATGAGATTATCAAAAGGATCTTCAACCTAGATCCTTTTAAATTA 4767
Db 718 TTAAGGATTTTGGTCAATGAGATTATCAAAAGGATCTTCAACCTAGATCCTTTTAAATTA 777
QY 4768 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCA 4827
Db 778 AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCA 837
QY 4828 ATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGCTTATCCATAGTTGC 4887
Db 838 ATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGCTTATCCATAGTTGC 897
QY 4888 CTGACTCCCGTGTGTAGATACTAGATACGGGGGCTTACCATCTGGCCCGAGTGC 4947
Db 898 CTGACTCCCGTGTGTAGATACTAGATACGGGGGCTTACCATCTGGCCCGAGTGC 957
QY 4948 TGAATGATACCGGAGACCCAGCTCACCAGGCTCCAGATTTATCAGCAATAAACAGGCC 5007
Db 958 TGAATGATACCGGAGACCCAGCTCACCAGGCTCCAGATTTATCAGCAATAAACAGGCC 1017
QY 5008 AGCGGAAGGGCGAGCGCAGAGTGTCT 5038
Db 1018 AGCGGAAGGGCGAGCGCAGAGTGTCT 1048
```

RESULT 3

```
BM438846/c 1013 bp mRNA linear EST 31-JAN-2002
LOCUS Iplvr00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
DEFINITION sequence.
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE Ictalurus punctatus (channel catfish)
ORGANISM Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 1013)
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
Transcriptome of channel catfish (Ictalurus punctatus) : initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
```

```
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..1013
/organism="Ictalurus punctatus"
/mol_type="mRNA"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"
ORIGIN
Query Match 16.8%; Score 986.4; DB 4; Length 1013;
Best Local Similarity 99.5%; Pred. No. 9.2e-289;
Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;
QY 4625 ATTACCGCGAGAAAAAGGATCTCAAGAGATCTTCAAGAGATCTTTTGAATCTTTTCTACGGGTCTGAC 4684
Db 1013 ATTACCGCGAGAAAAAGGATCTCAAGAGATCTTCAAGAGATCTTTTGAATCTTTTCTACGGGTCTGAC 954
QY 4685 GCTCAGTGGAAACGAAACCTCACGTTAAGGATTTTGGTCAAGATTTATCAAAAGGATC 4744
Db 953 GCTCAGTGGAAACGAAACCTCACGTTAAGGATTTTGGTCAAGATTTATCAAAAGGATC 894
QY 4745 TTCACCTAGATCTTTTAAATTAATTAAGTGTAAATCAATCTAAAGTATATATGAG 4804
Db 893 TTCACCTAGATCTTTTAAATTAATTAAGTGTAAATCAATCTAAAGTATATATGAG 834
QY 4805 TAACTTGGTCTGACAGTGTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 4864
Db 833 TAACTTGGTCTGACAGTGTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT 774
QY 4865 CTATTTTCGTTTCATCCATAGTTGCTGACCTCCCGTCTGTGTAGATAACTACGATACGGGAG 4924
Db 773 CTATTTTCGTTTCATCCATAGTTGCTGACCTCCCGTCTGTGTAGATAACTACGATACGGGAG 714
QY 4925 GGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGTCTCA 4984
Db 713 GGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGTCTCA 654
QY 4985 GATTTATCAGCAATAAACCCAGCCGAGGCGGCGAGAGTGTCTCTGCAACT 5044
Db 653 GATTTATCAGCAATAAACCCAGCCGAGGCGGCGAGAGTGTCTCTGCAACT 594
QY 5045 TTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAGTGTGCGCCA 5104
Db 593 TTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAGTGTGCGCCA 534
QY 5105 GTTAATAGTTTGGCGAAACGTTGTTGCCATTTGCTACAGGCATCTGTGTGTCAAGTGTGCG 5164
Db 533 GTTAATAGTTTGGCGAAACGTTGTTGCCATTTGCTACAGGCATCTGTGTGTCAAGTGTGCG 474
QY 5165 TTTGGTATGGCTTCATTCAGTCCGCTTCCCAACGATCAAGGCGAGTACA-TGATCCCC 5223
Db 473 TTTGGTATGGCTTCATTCAGTCCGCTTCCCAACGATCAAGGCGAGTACA-TGATCCCC 414
QY 5224 CATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCTCCGATCGTTGTTCAGAAAGTAGTT 5283
Db 413 CATGTTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCTCCGATCGTTGTTCAGAAAGTAGTT 355
QY 5284 GCGCGCAGTGTATCACTCATGTTATGACAGCAGTGCATAATTTCTTACTGTCTATGCC 5343
Db 354 GCGCGCAGTGTATCACTCATGTTATGACAGCAGTGCATAATTTCTTACTGTCTATGCC 295
QY 5344 ATCCGTAAGATGCTTTTCTGTGATCTGGTGTACTCAACCAAGTCAATCTGAGAAATAGTG 5403
Db 294 ATCCGTAAGATGCTTTTCTGTGATCTGGTGTACTCAACCAAGTCAATCTGAGAAATAGTG 235
QY 5404 TATGCGGCGACCGAGTGTCTTTCGCGCGGTCAATACGGGATAATACCGCGCCACATAG 5463
Db 234 TATGCGGCGACCGAGTGTCTTTCGCGCGGTCAATACGGGATAATACCGCGCCACATAG 175
QY 5464 CAGAACTTTAAAGTGTCTCATCATTTGGAAAAAGTCTTTCGGGGCGAAAAAATCTCTCAAGGAT 5523
|||||
```

Db	174	CAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTCGGGGCGAAAACTCTCAAGGAT	115
QY	5524	CTTTACCGCTGTTGAGATTCAGATTGTAACCACTCTGTGACCACTGATCTTTCAGC	5583
Db	114	CTTTACCGCTGTTGAGATTCAGATTGTAACCACTCTGTGACCACTGATCTTTCAGC	55
QY	5584	ATCTTTTACTTTCACCAAGGCTTCTCGGTGAGCAAAACAGGAGGCAAAATGC	5637
Db	54	ATCTTTTAACTTTACCAAGGCTCTCTGGGTGAGCAAAACAGGAGGCAAAATGC	1
RESULT 4			
AJ281480			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 999; Conservative			
QY	4023	AAAAGCCAGGAACCGTTAAAGAGCGCGTGTGCTGGCGTTTTTCCATAGGCTCCGCCCC	4082
Db	1	AAAAGCCAGGAACCGTTAAAGAGCGCGTGTGCTGGCGTTTTTCCATAGGCTCCGCCCC	60
QY	4083	CTGACGAGCATCACAAATTCAGCGCTCAAGTCAGAGGTGGCAACCCGACGAGCTAT	4142
Db	61	CTGACGAGCATCACAAATTCAGCGCTCAAGTCAGAGGTGGCAACCCGACGAGCTAT	120
QY	4143	AAAAGATACCAAGGCTTTCCCGCTTCCCGTGGAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGC	4202
Db	121	AAAAGATACCAAGGCTTTCCCGCTTCCCGTGGAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGC	180

QY	4203	CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCTGGCGCTTTCTCAATGCT	4262
Db	181	CGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCTGGCGCTTTCTCAATGCT	240
QY	4263	CACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAAAGCTGGGCTGTGTGCAG	4322
Db	241	CACGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCCAAAGCTGGGCTGTGTGCAG	300
QY	4323	AACCCCGCTTCAGCCCGACCGCTCGCTTATCCGGTAACTATCGTTTGTAGTCCAACC	4382
Db	301	AACCCCGCTTCAGCCCGACCGCTCGCTTATCCGGTAACTATCGTTTGTAGTCCAACC	360
QY	4383	CGGTAAAGACAGCTTATCGCCACTGGCAGCAGCAGCTGGTAAACAGGATTAGCAGAGCGA	4442
Db	361	CGGTAAAGACAGCTTATCGCCACTGGCAGCAGCAGCTGGTAAACAGGATTAGCAGAGCGA	420
QY	4443	GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGCTACACTAGAA	4502
Db	421	GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACGCTACACTAGAA	480
QY	4503	GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA	4562
Db	481	GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA	540
QY	4563	GCTTTGATCCGGGAAAAAACCACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGC	4622
Db	541	GCTTTGATCCGGGAAAAAACCACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGC	600
QY	4623	AGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATCTTTCTACGGGTCTG	4682
Db	601	AGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATCTTTCTACGGGTCTG	660
QY	4683	ACGCTCAGTGGAAACGAAAACTCACGTTAAGGAGTTTGTCTATGAGATTATCAAAAAAG	4742
Db	661	ACGCTCAGTGGAAACGAAAACTCACGTTAAGGAGTTTGTCTATGAGATTATCAAAAAAG	720
QY	4743	TCTTACCTAGATCTCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATATG	4802
Db	721	TCTTACCTAGATCTCTTTTAAATTAAGATTAAGTTTAAATCAATCTAAAGTATATG	780
QY	4803	AGTAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGAGCAGCTATCTCAGCGATCT	4862
Db	781	AGTAACTTTGGTCTGACAGTTTACCAATGCTTAAATCAGTGGAGCAGCTATCTCAGCGATCT	839
QY	4863	GTCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTGGTGTAGATACTACGATACGGG	4922
Db	840	GTCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTGGTGTAGATACTACGATACGGG	898
QY	4923	AGGGCTTACCATCTGCGCCCGAGTGTGCATGATACCGGAGACCCAGCTCACCGGCTC	4982
Db	899	AGGGCTTACCATCTGCGCCCGAGTGTGCATGATACCGGAGACCCAGCTCACCGGCTC	955
QY	4983	CAGATTATTCAGCAATAAACACGAGCCAGCGGAGGCGGAGCGAGCAAGTG	5033
Db	956	CAGATTATTCAGCAATAAACACGAGCCAGCGGAGGCGGAGCGAGCAAGTG	1004
RESULT 5			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			


```

/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/note="Vector: pSport1_sfi; Site_1: SfiIa; Site_2: SfiIB"

ORIGIN
Query Match      14.9%; Score 876.4; DB 7; Length 902;
Best Local Similarity 99.7%; Pred. No. 3.7e-255;
Matches 899; Conservative 1; Indels 2; Gaps 2;

QY 4906 GATAACTACGATACGGAGGGCTTACCATCTGGGCCCCAGTGTCTGCAATGATACCGCAGA 4965
DB 901 GATTAACACGATACGGAGGGCTTACCATCTGGGCCCCAGTGTCTGCAATGATACCGCAGA 842

QY 4966 CCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACACGACCGCCGGAAGGCGCGAGCG 5025
DB 841 CCCACGCTCACCGGCTCCAGATTTTTCAGCAATTAACACGACCGCCGGAAGGCGCGAGCG 782

QY 5026 CAGAAGTGGTCTCCCACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGGAAGC 5085
DB 781 CAGAAGTGGTCTCCCACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGGAAGC 722

QY 5086 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGCACAAAGTGTGTCCTATGCTACAGGCAT 5145
DB 721 TAGAGTAAGTAGTTGCGCAGTTAATAGTTTGCAGAACGTTGTCCTATGCTACAGGCAT 663

QY 5146 CGTGGTGTACGCTCGTGTGGTGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAG 5205
DB 662 CGTGGTGTACGCTCGTGTGGTGTATGGCTTCATTCAGCTCCGGTTCCTCAACGATCAAG 603

QY 5206 GCGAGTTTACATGATCCCCATGTTGTCACAAAAGCGTTAGTCTTTCGTCCTCCGAT 5265
DB 602 GCGAGTTTACATGATCCCCATGTTGTCACAAAAGCGTTAGTCTTTCGTCCTCCGAT 544

QY 5266 CGTTGTCAGAAAGTAGTTGGCCGAGTGTTATCACTCATGTGTTATGGCAGCACTGCATAA 5325
DB 543 CGTTGTCAGAAAGTAGTTGGCCGAGTGTTATCACTCATGTGTTATGGCAGCACTGCATAA 484

QY 5326 TTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA 5385
DB 483 TTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA 424

QY 5386 GTCAATCTTGAGAAAGTAGTATGCGCGACCGAGTTGCTTCTTGGCCGCGTCAATACGGGA 5445
DB 423 GTCAATCTTGAGAAAGTAGTATGCGCGACCGAGTTGCTTCTTGGCCGCGTCAATACGGGA 364

QY 5446 TAATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCTCGG 5505
DB 363 TAATAACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTCTTCTCGG 304

QY 5506 GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTC 5565
DB 303 GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTC 244

QY 5566 ACCCAACTGATCTTTCAGCATCTTTTACTTTTACCACAGGCTTCTCGGTGAGCAAAAAAGG 5625
DB 243 ACCCAACTGATCTTTCAGCATCTTTTACTTTTACCACAGGCTTCTCGGTGAGCAAAAAAGG 184

QY 5626 AAGCAAAATGCCGAAAAAAGGAAATAAGGGGACACGGAATATGTTGAATACTCATACT 5685
DB 183 AAGCAAAATGCCGAAAAAAGGAAATAAGGGGACACGGAATATGTTGAATACTCATACT 124

QY 5686 CTTCCTTTTCAATATTAATTGAACATTTATCAGGTTATGTTCTCATGAGCGGATACAT 5745
DB 123 CTTCCTTTTCAATATTAATTGAACATTTATCAGGTTATGTTCTCATGAGCGGATACAT 64

QY 5746 ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCCGGCACATTTCCCGAAAAAGT 5805
DB 63 ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCCGGCACATTTCCCGAAAAAGT 4

QY 5806 GC 5807
DB 3 GC 2

```

RESULT 7	EG838279/c	EG838279	935 bp	mRNA	linear	EST 25-MAY-2001
LOCUS		Gc01.10e07_R Gc01_AAFc ECORC cold_stressed_Glycine clandestina				
DEFINITION		Glycine clandestina cDNA clone Gc01_10e07_mRNA sequence.				
ACCESSION		EG838279				
VERSION		EG838279.1	GI:14204601			
KEYWORDS		EST.				
SOURCE		Glycine clandestina				
ORGANISM		Glycine clandestina				
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
REFERENCE		1 (bases 1 to 935)				
AUTHORS		Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.				
TITLE		Expressed Sequence Tags from Cold-Stressed Glycine clandestina				
JOURNAL		Seedlings				
COMMENT		Unpublished (2001) Contact: Singh,J.A. Eastern Cereal and Oilseed Research Centre Agriculture and Agri-food Canada KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, Canada Tel: (613) 759-1662 Fax: (613) 759-1701 Email: singhja@agr.gc.ca.				
FEATURES		Location/Qualifiers				
source		1..935 /organism="Glycine clandestina" /mol_type="mRNA" /cultivar="1035" /db_xref="taxon:45687" /clone="Gc01.10e07" /tissue_type="Leaves, stem" /clone_lib="Gc01_AAFc_ECORC_cold_stressed_Glycine_clandestina"				
ORIGIN		/notes=Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Plants incubated at 2 degrees under 12 hours of light/day. Harvested after only 2-3 days of cold treatment. cDNA was prepared with the Uni-Zap cDNA kit from Stratagene. Eco RI adapters were linked followed by digest with Xho I/Eco RI and ligated to pBluescript."				
Query Match		14.9%	Score 875.8;	DB 4;	Length 935;	
Best Local Similarity		97.7%	Pred. No. 5.6e-255;			
Matches		882;	Conservative 14;	Mismatches 6;	Indels 1;	Gaps 1;
Qy	4921	GGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGC	4980			
Db	928	GGGGGGCTTACCATCTGCCCCAGKGTGCAATGATACCGCGAGMCCACGCTCMCCGC	869			
Qy	4981	TCAGATTTATCAGCAATAACACGACCGCGGAGCGCGAGCGAAGTGTGCTCTGC	5040			
Db	868	YCCAGATTTATCAGCAATAAMCCACGCGGAGGCGCGAGCGAAGTGTGCTCTGC	809			
Qy	5041	AACTTTATCCGCTCCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTC	5100			
Db	808	AWCTTTATCCGCTCCCATCCAGTCTATTAAATTTGTCGGGAAAGCTAGAGTAAGTAGTTC	749			
Qy	5101	GCCAGTTAATAGTTTGGCGAAAGTTGTCATTGTCATTGTCATTGTCATTGTCATTGTC	5160			
Db	748	GCCAGTTAATAGTTTGGCGAAAGTTGTCATTGTCATTGTCATTGTCATTGTCATTGTC	689			
Qy	5161	GTGTTTGGTATGGCTTCATTACGCTCCGGTCCCAACGATCAAGGCGAGTACATGATC	5220			
Db	688	GTGTTTGGTATGGCTTCATTACGCTCCGGTCCCAACGATCAAGGCGAGTACATGATC	629			
Qy	5221	CCCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCCGTCCTCCGATCGTTGTGCAGAGTAA	5280			

```
Db 628 CCCCATGTTGTGCAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTGTCAAGATAA 569
Qy 5281 GTT-GGCGGAGGTTATCACTCATGTTATGCGACACTGCATAATCTCTTACTGTCA 5339
Db 568 GTTGGGCGGAGGTTATCACTCATGTTATGCGACACTGCATAATCTCTTACTGTCA 509
Qy 5340 TGCCATCCGTAAGATGCTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTGAGAAT 5399
Db 508 TGCCATCCGTAAGATGCTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTGAGAAT 449
Qy 5400 AGTGTATGCGGCGACCGAGTGTCTTTCGCCGCGTCAATACGGGATAATACCGCGCCAC 5459
Db 448 AGTGTATGCGGCGACCGAGTGTCTTTCGCCGCGTCAATACGGGATAATACCGCGCCAC 389
Qy 5460 ATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAGTCTTTCGGGCGGAAACTCTCAA 5519
Db 388 ATAGCAGAACTTTAAAGTCTCATCATTTGGAAGAGTCTTTCGGGCGGAAACTCTCAA 329
Qy 5520 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCAACTGATCTT 5579
Db 328 GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCAACTGATCTT 269
Qy 5580 CAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCG 5639
Db 268 CAGCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCG 209
Qy 5640 CAAAAAGGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTCTCTTTTCAAT 5699
Db 208 CAAAAAGGGAATAAGGGGACACGGAATGTTGAATACTCATCTCTCTCTTTTCAAT 149
Qy 5700 ATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGGAATGATTT 5759
Db 148 ATTATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGGAATGATTT 89
Qy 5760 AGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTCT 5819
Db 88 AGAAAAATAAACAAATAGGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGAAATGT 29
Qy 5820 AAG 5822
Db 28 AAG 26

RESULT 8
AL042026/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

841 bp mRNA linear EST 04-SEP-2003
DKFP434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFP434E111 5', mRNA sequence.

AL042026
AL042026.1 GI:5421372
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 841)
Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Poustka, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFP434E111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```
FEATURES
source
Location/Qualifiers
1..841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP434E111"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 14.3%; Score 841; DB 1; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.3e-244;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4936 TGCCCCCAGTGTGCAATGATACCGCGAGACCCACGCTACCGGCTCCAGATTATCAGC 4995
Db 841 TGCCCCCAGTGTGCAATGATACCGCGAGACCCACGCTACCGGCTCCAGATTATCAGC 782
Qy 4996 AATAAACACGAGCCGAGAGGCGCGAGAGTGGTCTTGCACACTTTATCCGCTC 5055
Db 781 AATAAACACGAGCCGAGAGGCGCGAGAGTGGTCTTGCACACTTTATCCGCTC 722
Qy 5056 CATCCAGTCTATTAAATGTTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 5115
Db 721 CATCCAGTCTATTAAATGTTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT 662
Qy 5116 GGCACAGTGTGGCATTGCTACAGGCATCGTGGTGTCA CGCTCGTCTGTTGGTATGCG 5175
Db 661 GGCACAGTGTGGCATTGCTACAGGCATCGTGGTGTCA CGCTCGTCTGTTGGTATGCG 602
Qy 5176 TTCATTTCAGTCCGGTTCCTCCAGATCAAGGCGAGTTACATGATCCCCCACTGTTGTCAA 5235
Db 601 TTCATTTCAGTCCGGTTCCTCCAGATCAAGGCGAGTTACATGATCCCCCACTGTTGTCAA 542
Qy 5236 AAAAGCGGTAGTCTCTCGGTCTCCGATCGTGTGTGAGAAAGTAAAGTGGCCGAGTGT 5295
Db 541 AAAAGCGGTAGTCTCTCGGTCTCCGATCGTGTGTGAGAAAGTAAAGTGGCCGAGTGT 482
Qy 5296 ATCAGTATGTTATGCGAGCAGTGCATAAATCTCTTACTGTCTATGTCATGCGTAAAGATG 5355
Db 481 ATCAGTATGTTATGCGAGCAGTGCATAAATCTCTTACTGTCTATGTCATGCGTAAAGATG 422
Qy 5356 CTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTGAGATAGTATGCGGAC 5415
Db 421 CTTTCTGTGACTGGTGTGAGTACTCAACCAAGTCAATCTGAGATAGTATGCGGAC 362
Qy 5416 GAGTTGCTCTTGGCCGCGTCAATACGGGATATACCGGCGCACATAGCAGAACTTTAAA 5475
Db 361 GAGTTGCTCTTGGCCGCGTCAATACGGGATATACCGGCGCACATAGCAGAACTTTAAA 302
Qy 5476 AGTGTCTCATCTTGGAAAAAGTGTCTTTCGGGCGAAAACTCTCAAGGATCTTACCGTGT 5535
Db 301 AGTGTCTCATCTTGGAAAAAGTGTCTTTCGGGCGAAAACTCTCAAGGATCTTACCGTGT 242
Qy 5536 GAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTTCAGCATCTTTTACTTT 5595
Db 241 GAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTTCAGCATCTTTTACTTT 182
Qy 5596 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCGCAAAAAGGAATAAG 5655
Db 181 CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCGCAAAAAGGAATAAG 122
Qy 5656 GCGCACACGAAATGTTGAATACTCATCTCTTCTTTTCAATATTTATTTGAAGCAATTTA 5715
Db 121 GCGCACACGAAATGTTGAATACTCATCTCTTCTTTTCAATATTTATTTGAAGCAATTTA 62
Qy 5716 TCAGGGTTATTGTCTCATGAGCGGATACATATTTTGAATGATTTTGAATAAATAACAAAT 5775
Db 61 TCAGGGTTATTGTCTCATGAGCGGATACATATTTTGAATGATTTTGAATAAATAACAAAT 2
```



```
/sub_species="vulgare"  
/db_xref="GABI:946193"  
/db_xref="taxon:112509"  
/clone="MPMG2010B0311"  
/tissue_type="embryosac"  
/dev_stage="0-10 DAF (days after flowering)"  
/lab_host="E. coli, SCS-1/pSE111"  
/clone_lib="BES1824"  
/notes="Vector: pOE30NST (AF074376); Site 1: SalI; Site 2:  
NotI; 0-10 DAF (days after flowering), cDNA synthesis  
using pBluescript II XR cDNA-library construction kit  
(Stratagene) with an oligo(dT)-primer containing NotI  
restriction site and a SalI adapter (Invitrogen). The main  
library of 21500 clones was rearranged into the sublibrary  
BES 1824 containing 4100 putative expression clones. Note:  
Due to a cloning artefact caused by the kit, in most cases  
the SalI site is NOT present, as well as the SalI Adapter  
used for cloning. To excise the insert, restriction sites  
upstream SalI should be used (e.g. BamHI). Average insert  
size is 1 kb. Library generation and sequencing was  
granted in context of GABI; data are also accessible at  
https://gabi.rzpd.de"
```

ORIGIN

```
Query Match 13.9%; Score 817; DB 7; Length 865;  
Best Local Similarity 98.0%; Pred. No. 5.1e-237;  
Matches 847; Conservative 0; Mismatches 13; Indels 4; Gaps 2;  
  
QY 3905 TCCTGGCTCGCTCGCTGGCTGGCGGAGCGGTATCAGCTCACTCAAGGGCGTAATA 3964  
DB 1 TCCTGGCTCGCTCGCTGGCTGGCGGAGCGGTATCAGCTCACTCAAGGGCGTAATA 60  
  
QY 3965 CGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATGAGCAAAAGGCCAGCA 4024  
DB 61 CGGTTATCCACAGATCAGGGGATAACGAGGAAAGAACATGAGCAAAAGGCCAGCA 120  
  
QY 4025 AAGGCCAGGAACGTAAGGCGCGTTCCTGGCGTTTTCATAGGCTCCGCCCCCT 4084  
DB 121 AAGGCCAGGAACGTAAGGCGCGTTCCTGGCGTTTTCATAGGCTCCGCCCCCT 180  
  
QY 4085 GAGCAGATCACAATAATCAGCTCAAGTCAAGTGGGGAACCCGACAGACTATAA 4144  
DB 181 GAGCAGATCACAATAATCAGCTCAAGTCAAGTGGGGAACCCGACAGACTATAA 240  
  
QY 4145 AGATACAGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCCTGTTCGACCTGCGG 4204  
DB 241 AGATACAGCGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCCTGTTCGACCTGCGG 300  
  
QY 4205 CTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTCA 4264  
DB 301 CTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTCA 360  
  
QY 4265 CGCTGAGGTATCTCAGTTCGGTGTAGTGTTCGCTTCCAGCTGGGCTGTGTGCAGAA 4324  
DB 361 CGCTGAGGTATCTCAGTTCGGTGTAGTGTTCGCTTCCAGCTGGGCTGTGTGCAGAA 420  
  
QY 4325 CCCCCGTTTCAGCGCGGACCGCTCGCTTACGTAATCTCGTTCGAGTCCACCCG 4384  
DB 421 CCCCCGTTTCAGCGCGGACCGCTCGCTTACGTAATCTCGTTCGAGTCCACCCG 480  
  
QY 4385 GTAAGACACGACTATTCGCCACTGGCAGCAGCCACTGGTAACAGGATTCAGAGCGAGG 4444  
DB 481 GTAAGACACGACTATTCGCCACTGGCAGCAGCCACTGGTAACAGGATTCAGAGCGAGG 540  
  
QY 4445 TATGTAGCGGCTACAGAGTTCTTGAAGTGGTGGCTTAATACCGGCTACCTAGAGG 4504  
DB 541 TATGTAGCGGCTACAGAGTTCTTGAAGTGGTGGCTTAATACCGGCTACCTAGAGG 600  
  
QY 4505 ACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGC 4564  
DB 601 ACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGC 660  
  
QY 4565 TCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCAGAGCAGCAG 4624
```

```
DB 661 TCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGTTCAGAGCAGCAG 720  
QY 4625 ATTACCGCAGAAAAAGAGATCTCAAGAGATCCCTTGTGATCTTCTACGGGTCTGAC 4684  
DB 721 ATTACCGCAGAAAAAGAGATCTCAAGAGATCCCTTGTGATCTTCTACGGGTCTGAC 780  
QY 4685 GCTCAGTGGAAACCAAACTCACCTT--AAGGGATTTTGGTCAATGA--GATTATCAAAAAG 4740  
DB 781 GCTCAGTGGAAACCAAACTCACCTTAAACGGGATTTTGGTCAATGAATATCAAAAAG 840  
QY 4741 GATCTTCACTAGATCTTTTAA 4764  
DB 841 GATCTTCACTANATCTTTTAA 864  
  
RESULT 11  
BZ577702/c  
LOCUS  
DEFINITION  
1126 bp DNA linear GSS 17-DEC-2002  
msh2 5533.Y2 msh Pseudomonas aeruginosa genomic clone msh2_5533,  
genomic survey sequence.  
ACCESSION  
BZ577702.1 GI:27212763  
VERSION  
GSS.  
KEYWORDS  
Pseudomonas aeruginosa  
Pseudomonas aeruginosa  
ORGANISM  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.  
REFERENCE  
1 (bases 1 to 1126)  
AUTHORS  
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,  
Burns,J.L., Kaul,R. and Olsen,M.V.  
TITLE  
Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
JOURNAL  
J. Bacteriol. (2002) In press  
COMMENT  
Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
LOCATION/Qualifiers  
1..1126  
/organism="Pseudomonas aeruginosa"  
/mol_type="genomic DNA"  
/strain="WSH"  
/db_xref="taxon:287"  
/clone="msh2_5533"  
/clone_lib="msh"  
/notes="Environmental isolate. Whole genomic shotgun  
library."  
  
ORIGIN  
  
Query Match 13.8%; Score 814.4; DB 8; Length 1126;  
Best Local Similarity 90.4%; Pred. No. 3.5e-236;  
Matches 938; Conservative 0; Mismatches 91; Indels 9; Gaps 6;  
  
QY 4735 AAAAGGATCTTCCCTAGATCCCTTTTAAATTAATAAAGTAAATCAATCTAAAG 4794  
DB 1125 AAAAGGTTTCTTCCCTAGACCCCTT--AATATAAAGAAAGTTTAAATACATTTTAAAG 1069  
QY 4795 TATATATGATTAACCTTGGTCTGACAGTTACCACTTAATCAGTGAGGCACC-TATCT 4853  
DB 1068 --TATTGAGTAACTTGTTCGACAGTGTCAAAGTTATATCATGAAGCCCCCTTATCA 1011  
QY 4854 CAGCATCTCTCTATTTTCGTTTCATCC-ATAGTTGCTCTGACTCCCC-GTCTGTAGATAAC 4911  
DB 1010 AGCTATCTCTATTTTCGTTTAAACAATAGTTCCTGACTCCCGTCCGGTAGAAT 951  
QY 4912 TAGCATACGGAGGGCTTACCATCTCGGCCCACTGTGTGCAATGATACCGGAGACCCAGC 4971  
DB 950 TCCGATCCGGAGGGTTTCTTATTTGCCCCCAAGTAGTCAATGTAACCGGAGCCCCCG 891
```



```
QY 4972 CTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGAG 5031
Db 890 TTCCCGCGCTCCAATTTATCCCAATAATACCAAGCATCCGAAAGGGCCGAGCCCAAAAG 831
QY 5032 TGGTCTCTGCAAC-TTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAG 5090
Db 830 TGGTCTCTGCAACTTTTATCAGCTCTCATCAGTCTATTAATTGTTGCCGGGAAGCTAGAG 771
QY 5091 TAAAGTAGTTCGCCAGTAATAGTTTTCGCAACAGTGTGTTCCCATGCTCAGGCACTGTGG 5150
Db 770 TCAGTAGTTTCGCCAGTTAATAGTATGCGCAACAGTGTGTTCCCATGCTCAGGCACTGTGG 711
QY 5151 TGTACGCTCTGCTGTTGGTATGGCTTCAATCAGCTCCGGTTCCTCCCAACATCAAGCGAG 5210
Db 710 TGTACGCTGGTGGTATGGCTTCAATCAGCTCCGGTTCCTCCCAACATCAAGCGAG 651
QY 5211 TTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTTG 5270
Db 650 TTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTTG 591
QY 5271 TCAGAAGTAGTTCGCCAGTGTATATCACTCATGTTATGCGAGCACTGCATAATTCTC 5330
Db 590 TCATAAGTAGTTCGCCAGTGTATATCACTCATGTTATGCGAGCACTGCATAATTCTC 531
QY 5331 TTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGTGAGTACTCAACCAAGTCA 5390
Db 530 TTACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGATGCTGTGAGTACTCAACCAAGTCA 471
QY 5391 TCTGAGAATAGTGTATGCGCGGACCGAGTGTCTTTCGCCGCGCTCAATACGGGATAATA 5450
Db 470 TCTGAGAATAGTGTATGCGCGGACCGAGTGTCTTTCGCCGCGCTCAATACGGGATAATA 411
QY 5451 CCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCATCATTTGGAACAGTCTTCGCGGCGAA 5510
Db 410 CCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCATCATTTGGAACAGTCTTCGCGGCGAA 351
QY 5511 AACTCTCAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCTGTGACCCA 5570
Db 350 AACTCTCAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCTGTGACCCA 291
QY 5571 ACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTCCTCGGTGAGCAAAAACAGGAAGC 5630
Db 290 ACTGATCTTCAGCATCTTTTACTTTTCCACAGCGTTCCTCGGTGAGCAAAAACAGGAAGC 231
QY 5631 AAAATGCCCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATCATCTACTCTTCC 5690
Db 230 AAAATGCCCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATCATCTACTCTTCC 171
QY 5691 TTTTTCATATTTGAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTG 5750
Db 170 TTTTTCATATTTGAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTG 111
QY 5751 AATGATTTTAGAAAAATA 5768
Db 110 AATGATTTTAGAAAAATA 93
```

RESULT 12

CN823189/c

LOCUS

CN823189 856 bp mRNA linear EST 02-JUN-2004
Oa_splbn_06L10_M13reverse Sheep spleen\brain pSport1 library Ovis

arles cDNA clone Oa_splbn_06L10 5', mRNA sequence.

ACCESSION

CN823189

VERSION

CN823189.1

KEYWORDS

EST.

SOURCE

Ovis aries (sheep)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Caprinae; Ovis.

REFERENCE

1 (bases 1 to 856)

AUTHORS

Gosner, A. and Hopkins, J.

TITLE

Ovine spleen\brain cDNA library

Unpublished (2004)

Contact: J Hopkins

Veterinary Biomedical Sciences

University of Edinburgh

Summerhall Square, Edinburgh, EH9 1QH.

Email: j.hopkins@ed.ac.uk

Plate: 06 row: L column: 10

Seq primer: M13reverse

High quality sequence start: 6

High quality sequence stop: 548.

FEATURES

Location/Qualifiers

1..856

/organism="Ovis aries"

/mol_type="mRNA"

/db_xref="taxon:9940"

/clone_lib="Oa_splbn_06L10"

/clone_lib="Sheep spleen\brain pSport1 library"

/note="Vector: pSport1"

ORIGIN

Query Match

Best Local Similarity 13.8%; Score 809.4; DB 7; Length 856;

Matches 830; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 4964 GACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAG 5023

Db 856 GACCCAGCTCACCGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAG 798

QY 5024 CCGAGAAGTGGTCTTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAA 5083

Db 797 CCGAGAAGTGGTCTTCGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAA 738

QY 5084 GCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGCACAAAGTGTGTCCTATGCTACAGGC 5143

Db 737 GCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGCACAAAGTGTGTCCTATGCTACAGGC 678

QY 5144 ATCGTGGTGTACGCTCGTTCGTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACGATCA 5203

Db 677 ATCGTGGTGTACGCTCGTTCGTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACGATCA 618

QY 5204 AGCGAGTGTACATGATCCCATGTTGTGCAAAAAGCGTTAGCTTCCTCGTCTCCG 5263

Db 617 AGCGAGTGTACATGATCCCATGTTGTGCAAAAAGCGTTAGCTTCCTCGTCTCCG 558

QY 5264 ATCGTGTGCAGAAAGTAAAGTGGCCGAGTGTATCACTCATGCTTATGCGAGCACTGCAT 5323

Db 557 ATCGTGTGCAGAAAGTAAAGTGGCCGAGTGTATCACTCATGCTTATGCGAGCACTGCAT 498

QY 5324 AATTCTCTTACTGTATGCCATCCGTTAGATGCTTTCTGTGACTGGTACTCAACC 5383

Db 497 AATTCTCTTACTGTATGCCATCCGTTAGATGCTTTCTGTGACTGGTACTCAACC 438

QY 5384 AAGTCATCTGAGAATAGTATGCGGACCGAGTGTGCTTTGCCCGCGCTCAATACGG 5443

Db 437 AAGTCATCTGAGAATAGTATGCGGACCGAGTGTGCTTTGCCCGCGCTCAATACGG 378

QY 5444 GATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCTGAGGAAAAAGTCTTCG 5503

Db 377 GATAATACCGGCCACATAGCAGAACTTTTAAAGTGTCTCATCTGAGGAAAAAGTCTTCG 318

QY 5504 GGGCGAAAACCTCTCAAGGATCTTACGGCTGTGAGATCCAGTCCAGTGTAAACCCACTCGT 5563

Db 317 GGGCGAAAACCTCTCAAGGATCTTACGGCTGTGAGATCCAGTGTAAACCCACTCGT 258

QY 5564 GCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTCCTGGGTGAGCAAAAAACA 5623

Db 257 GCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTCCTGGGTGAGCAAAAAACA 198

QY 5624 GGAAAGCAAAATGCGCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACTCATA 5683

Db 197 GGAAAGCAAAATGCGCGCAAAAAGGAATAAGGGCGACACGGAAATGTTGAATACTCATA 138

```
QY 5684 CTCTTCTTTTCAATATATTAAGCATTTATCAGGGTTATGTCTCATGACGGGATAC 5743
Db 137 CTCTTCTTTTCAATATATTAAGCATTTATCAGGGTTATGTCTCATGACGGGATAC 78
QY 5744 ATATTTGAATGTTTATAGAAAATAAACAATAGGGGTTCCGGCGACATTTCCCGCAAAA 5803
Db 77 ATATTTGAATGTTTATAGAAAATAAACAATAGGGGTTTCGACCCGGGAATCCGGACCG 18
QY 5804 GTGCCAC 5810
Db 17 GTACCCG 11

RESULT 13
LOCUS CD649375/c
DEFINITION Crassostrea virginica 819 bp mRNA linear EST 18-JUN-2003
5', mRNA sequence.
ACCESSION CD649375
VERSION CD649375.1 GI:31906346
SOURCE Crassostrea virginica (eastern oyster)
ORGANISM Crassostrea virginica
REFERENCE Crassostrea virginica; Bivalvia; Pteriomorpha; Ostreoida;
AUTHORS Eukaryota; Metazoa; Mollusca; Crassostrea.
TITLE 1 (bases 1 to 819)
DIFFERENTIALLY EXPRESSED OYSTER (Crassostrea virginica) genes after
exposure to mercury
JOURNAL Unpublished (2003)
COMMENT Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.

FEATURES
source Location/Qualifiers
1..819
/organism="Crassostrea virginica"
/mol_type="mRNA"
/db_xref="taxon:6565"
/clone_lib="Crassostrea virginica Gonad"
/notes="Organ: Gonad; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

ORIGIN
Query Match 13.7%; Score 808; DB 6; Length 819;
Best Local Similarity 99.9%; Pred. No. 2.8e-234;
Matches 819; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4818 ACAGTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTCTTATTTCTGTTCAAT 4877
Db 819 ACAGTTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTCTTATTTCTGTTCAAT 760
QY 4878 CCATAGTTGCTGACTCCCGTCTGTAGATACTAGATACGGAGGGCTTACCATCTG 4937
Db 759 CCATAGTTGCTGACTCCCGTCTGTAGATACTAGATACGGAGGGCTTACCATCTG 700
QY 4938 GCCCCAGTCTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTATACGAA 4997
Db 699 GCCCCAGTCTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGATTATACGAA 640
QY 4998 TAAACACGACCGGAGGCGGCGAGAGAGTGTCTCTGCAACTTTATTCGCGCTCCA 5057
Db 639 TAAACACGACCGGAGGCGGCGAGAGTGTCTCTGCAACTTTATTCGCGCTCCA 580
QY 5058 TCCAGTCTATTAATTTGTCGGGGAAGCTAGATAGTATGTTGCGCAGTTAATAGTTTC 5117
Db 579 TCCAGTCTATTAATTTGTCGGGGAAGCTAGATAGTATGTTGCGCAGTTAATAGTTTC 520
```

```
QY 5118 GCAACGCTTGTTCGCAATTCGTACAGGCATCGTGGTGTCAAGCTCGTCTGTTGGTATGCTT 5177
Db 519 GCACGCTTGTTCGCAATTCGTACAGGCATCGTGGTGTCAAGCTCGTCTGTTGGTATGCTT 460
QY 5178 CATTCAGCTCCGGTTCCTCCAAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAA 5237
Db 459 CATTCAGCTCCGGTTCCTCCAAACGATCAAGCGAGTTACATGATCCCCCATGTTGTGCAAAA 400
QY 5238 AAGCGGTTAGCTCCTTCCGTCCTCCGATCGTGTCTCAGAGTAGTGTGGCGCGAGTGTAT 5297
Db 399 AAGCGGTTAGCT-CITCGGTCTCTCCGATCGTGTCTCAGAGTAGTGTGGCGCGAGTGTAT 341
QY 5298 CACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTCTGTCATGCCATCCGTAAGATGCT 5357
Db 340 CACTCATGTTATGGCAGCACTGCATAATTCCTTACTGTCTGTCATGCCATCCGTAAGATGCT 281
QY 5358 TTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAAATAGTGTATCGCGCGACCGA 5417
Db 280 TTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGAAATAGTGTATCGCGCGACCGA 221
QY 5418 GTTGCTCTTGGCCGGCGTCAATACGCGGATAATACCGGCCACATAGCAGAACTTTAAAAG 5477
Db 220 GTTGCTCTTGGCCGGCGTCAATACGCGGATAATACCGGCCACATAGCAGAACTTTAAAAG 161
QY 5478 TGTCTCATCTATGGAAAAACGTTCTTCCGGGCGCAAAACTCTCAAGGATCTTTACCGCTGTGA 5537
Db 160 TGTCTCATCTATGGAAAAACGTTCTTCCGGGCGCAAAACTCTCAAGGATCTTTACCGCTGTGA 101
QY 5538 GATCCAGTTTCGATGTAAACCCACTCGTGTGACCCAACTGTATTCAGCATCTTTTACTTTCA 5597
Db 100 GATCCAGTTTCGATGTAAACCCACTCGTGTGACCCAACTGTATTCAGCATCTTTTACTTTCA 41
QY 5598 CCAGCGTTTCTCGGTGAGCAAAAACAGGAGCGCAAAATGC 5637
Db 40 CCAGCGTTTCTCGGTGAGCAAAAACAGGAGCGCAAAATGC 1

RESULT 14
LOCUS BZ576726/c
DEFINITION msh2 5071.y2 msh Pseudomonas aeruginosa genomic clone msh2_5071,
genomic survey sequence.
ACCESSION BZ576726
VERSION BZ576726.1 GI:27211787
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,W.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source Location/Qualifiers
1..1011
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2 5071"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun"
```

ORIGIN		Library."	
Query Match		13.6%; Score 797.6; DB 8; Length 1011;	
Best Local Similarity		94.0%; Pred. No. 4.6e-231;	
Matches 871; Conservative		0; Mismatches 52; Indels 4; Gaps 4;	
QY	4768	AAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGTCTGACAGTTACCA	4827
DB	993	AAAAATGAATTTTAAATCCCTCTCTAAAGATAAATTAGTAAACTTGTCTGCCAGGTTC-	935
QY	4828	ATGCTTAAATCAGTGGAGCCACTATCTCAGCGATCT-GTCTATTTCGTTTCATCCATAGTTG	4886
DB	934	ATGCTTAAATAGTAAGAGCCCATCTCAGCATCCGGTCTATTTCGGTCATCCATAGTTG	875
QY	4887	CTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGCTTTACCATCTGGCCCCAGTG	4946
DB	874	CCTGACTCCCGTCGTGTAGATAATTACCATACGGGAAGGTTTACCATCTGCCCAAGTC	815
QY	4947	CTGCAATGATACCGCAGAGCCACCGCT-CACCGCTCCAGATTTATCAGCAATAAACCCAG	5005
DB	814	TTGCAATAATACCCCGAGACCCACGCTCCACGGGCTCCAGATTTATCAGCAATAAACCCAG	755
QY	5006	CCAGCCGGAAGGCCCGAGCGCAGAAGTGTCTCTCAACTTTATCGCCCTCCATCCAGTCT	5065
DB	754	CAAGCCGAAGGGCCGAGCCCAAGTGTGCTCTGCAACTTTATCCGCTCCATCCAGTCT	695
QY	5066	ATTAATTGTTGCGGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAAGTTTGGCGCAACGTT	5125
DB	694	ATTAATTGTTG-CGGNAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTGCGCAACGTT	636
QY	5126	GTTCGCAATTGCTACAGGCATCTGTGTGTACGCTCGTCTGTTGGTATGCTTCATTCACG	5185
DB	635	GTTCGCAATTGCTACAGGCATCTGTGTGTACGCTCGTCTGTTGGTATGCTTCATTCACG	576
QY	5186	TCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTT	5245
DB	575	TCGGTTCCTCAACGATCAAGGCGAGTTACATGATCCNCCATGTTGTGCAAAAAGCGGTT	516
QY	5246	AGTCTCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCATCTCATG	5305
DB	515	AGTCTCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTATCATCTCATG	456
QY	5306	GTATGCGAGCACTGATTAATCTCTTACTGTATGTCATGCCATCCGTAAGATGCTTTTCTGTG	5365
DB	455	GTATGCGAGCACTGATTAATCTCTTACTGTATGTCATGCCATCCGTAAGATGCTTTTCTGTG	396
QY	5366	ACTGGTGAGTACTCAACCAAGTCATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCT	5425
DB	395	ACTGGTGAGTACTCAACCAAGTCATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCT	336
QY	5426	TGCCCCGCGTCAATACGGGATAATACCGGCCCATAGCAGAACTTTAAAAGTGTCTCATC	5485
DB	335	TGCCCCGCGTCAATACGGGATAATACCGGCCCATAGCAGAACTTTAAAAGTGTCTCATC	276
QY	5486	ATTGGAATCGTTCTTCGGGGGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGT	5545
DB	275	ATTGGAATCGTTCTTCGGGGGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGT	216
QY	5546	TCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTT	5605
DB	215	TCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGTT	156
QY	5606	TTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAGGGAATAAGGGGCACACGG	5665
DB	155	TTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAGGGAATAAGGGGCACACGG	96
QY	5666	AAATGTTGAATCTACTACTCTTCCTT	5692
DB	95	AAATGTTGAATCTACTACTCTTCCTT	69

[illegible]

Search completed: April 30, 2005, 10:14:10
Job time : 16339 secs

This Page Blank (uspto)